A Non-commutative Version of the Coupling from the Past Algorithm

Vom Fachbereich Mathematik
der Technischen Universität Darmstadt
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)
genehmigte Dissertation

von

M.phil. Nadiem Sissouno

aus Groß-Gerau

Referent:	Prof. Dr. B. Kümmerer
Koreferent:	Prof. Dr. H. Maassen
Tag der Einreichung:	10.12.2010
Tag der mündlichen Prüfung:	11.02.2011

Darmstadt 2011

D 17

Acknowledgements

Burkhard Kümmerer's research and mathematical ideas have provided fertile soil for my own ideas to grow. He always encouraged green shoots of discussion and creative thought throughout the course of this work. Burkhard is a mathematician who understands that life is not mathematics alone, and further understands that this can be more cause for concern than any false mathematical constructs or proofs.

I am similarly grateful to Hans Maassen for acting as co-referee. His efforts have made it possible for me to complete the body of this work in a form that is agreeable to all.

I also thank my predecessors from Burkhards group, whose work was a valuable source of reference, as well as those I have had the pleasure to work with. In particular I acknowledge my recent colleagues: Florian Haag, Lisa Steiner, Andreas Gärtner, Walter Reusswig and Kay Schwieger. I am grateful for their humour, friendship and numerous discussions that have assisted in formulating this thesis.

Finally, I thank Sonja Steberl for her support and for standing by me patiently all these years.

CONTENTS

Introduction 1

1 General Notation 9

1.1 A Few Facts about Operator Algebras 10

1.2 Non-commutative Probability Theory 12

1.2.1 Non-commutative Probability Space 12

1.2.2 Product Spaces 12

1.2.3 Morphisms 13

1.2.4 Conditional Expectation 16

1.2.5 Random Variables 17

1.2.6 Stochastic Processes 17

1.2.7 Markov Processes 19

1.2.8 Transition Operators 19

1.2.9 One-sided Commutative Coupling 21

1.2.10 One-sided Non-commutative Coupling 22

1.2.11 Two-sided Commutative Coupling 24

1.2.12 Two-sided Non-commutative Coupling 25

1.2.13 Micromaser 27

1.2.14 Scattering Theory 31

2 Graphs and Stochastic Matrices 35

2.1 Graphs . . . 35
2.1.1 Synchronising Words and Asymptotic Completeness . . . 37
2.2 Transition Matrices and Perron-Frobenius Theorem . . . 40
2.2.1 Irreducible and Aperiodic Markov Chains . . . 40
2.2.2 Perron-Frobenius Theorem . . . 42
2.3 Transition Operators and Perron-Frobenius Theorem . . . 43
2.3.1 Non-commutative Irreducible and Aperiodic Operators . . . 43
2.3.2 Irreducible Operators . . . 44
2.3.3 Non-commutative Perron-Frobenius Theorem . . . 48

3 Coupling from the Past 51
3.1 Composition into the Past . . . 53
3.2 Semigroups of Functions . . . 61
3.2.1 Semigroup and Update Function . . . 66
3.2.2 Global Couplings and Synchronising Words . . . 67
3.3 Résumé . . . 68

4 Synchronising Convex Decompositions of Operators 71
4.1 Synchronisable Operators . . . 71
4.2 Synchronisable Operators on M_n . . . 79
4.2.1 Arveson's Theorem . . . 80
4.2.2 Full Concrete Representation . . . 82
4.3 Synchronisable Operators on Finite Algebras . . . 87
4.4 Asymptotic Completeness and Synchronisation . . . 90
4.5 Perspectives . . . 95

5 Examples 97
5.1 Numerical Approach . . . 97
5.2 First Example . . . 101
5.2.1 Simulation . . . 102
5.3 Truncated Micromaser . . . 103
5.4 Second Example . . . 106
5.4.1 Simulation . . . 108
5.5 Third Example . . . 109

5.5.1 Simulation 111
5.6 Résumé 111

Glossary 115

Index 117

Bibliography 119

Introduction

There are numerous examples in computational areas of sciences, mentioning only physics, biology and linguistics, where algorithms are employed to produce random samples distributed according to a probability distribution which is either unknown or difficult to access. These techniques go by the name Markov Chain Monte Carlo methods. Usually they are applied where models are complex and multivariate and the distribution is not a product measure. The most famous have arisen in statistical physics. Ensembles of particles whose idealised interaction can be described by a Markov chain are analysed with regard to configurations, once they have reached an equilibrium, or phase transitions conditional to certain physical parameters. The application to the so called Ising model appears as an example in almost every modern textbook that covers the subject of Markov chains and sampling. More recent applications are to image analysis, compression and restoration, typically problems in computational engineering. The most popular sampling algorithms are the Metropolis-Hastings algorithm and the Gibbs sampler. Their disadvantage is that the samples are only approximately distributed like the required distribution; longer running time results in better approximation. Hence, the decision on the so called 'burn-in period' is left to the user. In the late 1980's variants of Markov Chain Monte Carlo methods, called exact or perfect, have been developed. Their major advantage is that they reach the target distribution in finite time almost surely and then stop automatically. In 1995 J. D. Propp and D. B. Wilson [PW96] published the so called Coupling from the Past algorithm, which drew the most attention among the exact sampling algorithms and initiated further investigation into this same subject. The term coupling in the above name refers to a useful method in stochastics, where two measures are analysed by considering another measure on a probability space whose marginal distributions coincide with one of those measures, respectively. As couplings are easily formulated for sequences of measures also, they may be used to study asymptotic behaviour of Markov chains with the same transition probabilities, but with different initial distribution. The sequence may be regarded as steps of an evolution into the future, in which the initial distribution marks fixed starting points at time $t = 0$. Descriptively, coupling produces sample paths from many initial distributions simultaneously. By choosing time $t = 0$ to be the target time and taking steps from the past towards $t = 0$, instead of assigning the starting time to $t = 0$ and proceeding from there, J. D. Propp and D. B. Wilson modified the classical coupling scheme and added a precise terminating condition.

Markov chains with given transition probabilities may usefully be represented as graphs. Consider the set of states as set of vertices and then create edges from transitions with positive probability. If the matrix of transition probabilities is aperiodic and irreducible, there exist representations as graphs, whose edges can be labelled, such that the labels determine the outgoing edges of every state uniquely. This class of graphs is sometimes called road-coloured and the graphs are called c-graphs. In addition, under the above assumption c-graphs can always be found, such that there exist finite sequences of labels that direct to a single vertex, irrespective of which vertex has been chosen initially. These are called synchronising words and have some relevance in coding theory or the theory of automata. We will develop a perspective on the Coupling from the Past algorithm, in which synchronising words take a key position; as we will see the termination of the algorithm coincides with the occurrence of a synchronising word.

Only recently R. Gohm, B. Kümmerer and T. Lang [GKL06] uncovered an interesting connection between synchronising words and asymptotic completeness. B. Kümmerer and H. Maassen introduced this asymptotic property to non-commutative probability theory. By dualisation c-graphs can be related to certain non-commutative random variables. One of these random variables is asymptotically complete if and only if the corresponding c-graph possesses a synchronising word. Initially, this equivalence has risen the question, whether it might be possible to formulate a non-commutative version of the Coupling of the Past algorithm, and motivated this research. Once started, however, asymptotic completeness faded to the background briefly. Our findings related to synchronising words and the Coupling to the Past algorithm led us to the definition of a so called synchronisable operator, a unital, completely positive operator on a C^*-algebra that features a convex decomposition into certain unital, completely positive operators. In non-commutative probability theory unital, completely positive operators play the part of transition matrices. With the classical algorithm being applicable to aperiodic, irreducible Markov chains, we turned to synchronisable operators and analysed their ergodic properties. On finite dimensional von Neumann algebras we have discovered that a unital, completely positive operator T possesses a synchronisable power T^n if and only if it is aperiodic and irreducible in the sense introduced in [FB09]. In this context the question concerning the relation to asymptotic completeness reappeared. For now it is only answered incompletely by proving that the transition operator of an asymptotically complete random variable has a synchronisable power. In the classical case of finite graphs with synchronising words and the corresponding asymptotically complete random variable, however, we obtain equivalence.

In **Chapter 1** we provide fundamental background to non-commutative probability theory. Basic concepts such as probability spaces and random variables are defined, as well as stochastic processes. In particular, Markov processes and transition operators, inextricably linked with those, are introduced. One way of constructing Markov processes is a coupling scheme on tensor product spaces by B. Kümmerer. Together with H. Maassen, he has also developed a scattering theory for those processes, which we present at the end of this chapter.
Chapter 2 is a miscellany, containing a section on graphs with a focus on graphs with coloured edges. The results of R. Gohm, B. Kümmerer and T. Lang that relate synchronising words and asymptotic completeness are discussed here. Although the following change to stochastic matrices, their classification and ergodic properties may seem sudden, it is not unconnected. As already mentioned, transition matrices may be presented by graphs.
The Coupling from the Past algorithm is introduced in **Chapter 3**. We develop a perspective that considers semigroups of functions on the state space, and bring out the relation to synchronising words.
With **Chapter 4** we leave the classical probability theory. After introducing synchronisable operators, we discuss their ergodic properties and show how their invariant state may be represented by an integral in state space with respect to the image measure of the backwards composition process. We continue with an analysis of synchronisable operators on finite dimensional von Neumann algebras. Considerations with regard to the connection of asymptotic completeness and

synchronisable operators then conclude this chapter.
In **Chapter 5** we suggest an approach to use synchronisable operators to produce random samples of states or approximate invariant states by means of limit theorems on Banach spaces. The larger part of this chapter is spent demonstrating findings of previous chapters.

Übersetzung der englischen Einleitung

Anhand zahlreicher Anwendungen läßt sich die Bedeutung algorithmischer Verfahren zur Erzeugung von zufälligen Strichproben, die eine unbekannte oder nur schwer zu erzeugende Verteilung besitzen, erkennen. In vielen Wissenschaften mit numerischen Problemstellungen - beispielsweise Physik, Biologie oder Linguistik - lassen sich Beispiele finden. Diese algoritmischen Methoden heißen Markov-Chain-Monte-Carlo Verfahren. Üblicherweise finden sie dann Anwendung, wenn die Modelle komplex und mulitvariat und die Verteilung kein Produktmaß ist. Die bekanntesten Beispiele entstammen der statistischen Physik. Hier werden Ensembles von Teilchen untersucht, deren idealisierte Wechselwirkungen durch Markov Ketten beschrieben werden kann. Zentrale Fragen sind hierbei ihre Konfiguration im Gleichgewicht oder Phasenübergänge in Abhängigkeit bestimmter physikalischer Parameter. Beinahe jedes moderne Lehrbuch, das Markov Ketten und Simulation behandelt, führt das Ising Modell als Beispiel an. Aktueller sind Anwendungen in Bildanalyse, -kompression und -wiederherstellung. Die bekanntesten Algorithmen sind der Metropolis-Hastings Algorithmus und der Gibbs Sampler. Der Nachteil dieser Verfahren ist allerdings, daß die erzeugten Stichproben nur annähernd die gewünschte Verteilung besitzen. Je länger der Algorithmus läuft, desto besser wird die Approximation. Die Entscheidung über die sogenannte "burn-in period" muß der Anwender treffen. Ende der 1980er Jahre wurden neue Markov-Chain-Monte-Carlo Verfahren entwickelt, die exakt oder perfekt genannt werden. Ihr wesentlicher Vorteil liegt darin begründet, daß sie die Zielverteilung fast sicher in endlicher Zeit erreichen und dann selbständig stoppen. 1995 veröffentlichten J. D. Propp und D. B.Wilson [PW96] den so genannten "Coupling from the Past" Algorithmus. Er ist unter allen exakten Algorithmen der bekannteste und regte zu intensiver Forschung auf diesem Gebiet an. Der Namensbestandteil "coupling" bezieht sich auf eine wichtige Methode in der Stochastik, die es erlaubt zwei Maße zu untersuchen, indem ein Maß auf einem Produktwahrscheinlichkeitsraum betrachtet wird, dessen Marginalverteilungen den beiden ursprünglichen Maßen entsprechen. Couplings lassen sich auch für Folgen von Maßen definieren, was es erlaubt, das asymptotische Verhalten von Markov Ketten zu untersuchen, die die selben Übergangswahrscheinlichkei-

ten, aber unterschiedliche Anfangsverteilungen besitzen. Die Anfangsverteilungen bestimmen die Startpunkte zur Zeit $t = 0$ und die einzelnen Glieder der Folgen Entwicklungsschritte in die Zukunft. Anschaulich erzeugt ein coupling gleichzeitig Trajektorien aus vielen Anfangsverteilungen. Indem J. D. Propp und D. B. Wilson $t = 0$ statt zur Start- zur Zielzeit erklärten und die Startzeit schrittweise weiter in die Vergangenheit verschoben, modifizierten sie das klassische coupling-Verfahren und schufen eine präzise Abruchbedingung.
Markov Ketten mit Übergangsmatrizen lassen sich mit Hilfe von Graphen darstellen. Die Menge der Zustände wird zur Menge der Ecken und Übergänge mit positiven Übergangswahrscheinlichkeiten erzeugen Kanten. Falls die Übergangsmatrix aperiodisch, irreduzibel ist, gibt es mindestens einen Graph, dessen Kanten so gefärbt werden können, daß die Farben die auslaufenden Kanten jeder einzelnen Ecke eindeutig bestimmen. Solche Graphen werden straßengefärbt genannt und die Graphen selbst als c-Graph bezeichnet. Darüber hinaus lassen sich unter den obigen Voraussetzungen immer c-Graphen finden, sodass eine endliche Folge von Farben existiert, die unabhängig von der Startecke zu einer festen Zielecke führt. Diese Folgen heißen synchronisierende Worte und sind sowohl in der Kodierungs- als auch in der Automatentheorie von Bedeutung. Wir werden eine Sichtweise auf den Coupling from the Past Algorithmus entwickeln, in der synchronisierende Worte ein Schlüßelrolle einnehmen. Wir werden sehen, dass der Algorithmus genau dann abbricht, wenn ein synchronisierendes Wort aufgetreten ist.
R. Gohm, B. Kümmerer and T. Lang [GKL06] konnten vor kurzem einen Zusammenhang zwischen synchronisierenden Worten und asymptotischer Vollständigkeit nachweisen. Letztere Eigenschaft hat B. Kümmerer in der nichtkommutativen Wahrscheinlichkeitstheorie etabliert. Durch Dualisieren lassen sich c-Graphen mit bestimmten nichtkommutativen Zufallsvariablen identifizieren. Eine solche Zufallsvariable ist genau dann asymptotisch vollständig, wenn der assozierte c-Graph ein synchronisierendes Wort besitzt. Diese Äquivalenz stieß die Frage an, ob es möglich ist eine nichtkommutative Version des Coupling from the Past Algorithmus zu formulieren, und motivierte unsere Forschung. Im Laufe der Untersuchung spielte die asymptotische Vollständigkeit zunächst jedoch eine untergeordnete Rolle. Die Erkenntnisse, die wir zu synchronisierenden Worten und dem Coupling from the Past Algorithmus sammelten, führten zu der Definition eines so genannten synchro-

nisierbaren Operators. Dabei handelt es sich um einen unitalen, vollständig positiven Operator auf einer C^*-Algebra, der eine Konvexzerlegung in bestimmte unitale, vollständig positiven Operatoren besitzt. Da der Algorithmus in der klassischen Wahrscheinlichkeitstheorie auf aperiodisch, irreduzible Markov Ketten anwendbar ist, konzentrierten wir uns auf die ergodischen Eigenschaften von synchronisierbaren Operatoren. Auf endlich dimensionalen Algebren konnten wir nachweisen, dass ein Operator T genau dann eine synchronisierbare Potenz besitzt, wenn er aperiodisch und irreduzible im Sinne von [FB09] ist. In diesem Zusammenhang tauchte die Frage nach dem Zusammenhang zu asymptotischer Vollständigkeit wieder auf. Die Antwort ist nur insoweit vollständig, als wir bisher zeigen konnten, dass der Übergangsoperator einer asymptotisch vollständigen Zufallsvariable eine synchronisierbare Potenz besitzt. Im klassischen Fall von endlichen Graphen mit synchronisierendem Wort und den assoziierten vollständig positiven Zufallsvariablen konnten wir allerdings eine Äquivalenz nachweisen.

In **Kapitel 1** stellen wir die notwendigen Grundlagen aus der nichtkommutativen Wahrscheinlichkeitstheorie. Wir definieren fundamentale Konzepte wie Wahrscheinlichkeitsraum, Zufallsvariable und stochastischer Prozess, insbesondere Markov Prozesse und Übergangsoperatoren. Ein Beispiel Markov Prozesse zu konstruieren, ist ein Kopplungsverfahren auf Tensorprodukt-Räumen von B. Kümmerer. Für diese Prozesse entwickelte er und H. Maassen eine Streutheorie, die am Ende dieses Kapitels vorstellen.
In Kapitel 2 konzentrieren wir uns auf gefärbte Graphen. Die Ergebnisse von R. Gohm, B. Kümmerer and T. Lang zu synchronisierenden Worten und asymptotischer Vollständigkeit werden hier diskutiert. Das restliche Kapitel befasst sich mit Übergangsmatrizen, ihrer Klassifikation und ihren ergodischen Eigenschaften.
In Kapitel 3 stellen wir den Coupling from the Past Algorithmus vor. Wir entwickeln eine Sichtweise die Halbgruppen von Funktionen auf dem Zustandsraum betrachtet und arbeiten den Zusammenhang zu synchronisierenden Worten heraus.
In Kapitel 4 verlassen wir die klassische Wahrscheinlichkeitstheorie. Nachdem wir synchronisierbare Operatoren eingeführt haben, widmen wir uns ihren ergodischen

Eigenschaften und zeigen, dass ihr invarianter Zustand als Integral über dem Zustandsraum bezüglich des Bildmaßes des Rückwärts-Verkettungsprozesses gegeben ist. Wir beenden dieses Kapitel mit Überlegungen zu asymptotischer Vollständigkeit und synchronisierenden Operatoren.

In Kapitel 5 stellen wir eine mögliche Herangehensweise vor, um zufällige Stichproben von Zuständen zu erzeugen oder den invarianten Zustand mit Hilfe von Grenzwertsätzen auf Banachräumen zu nähern. Im größten Teil dieses Kapitels demonstrieren wir unsere Resultate anhand von Beispielen.

CHAPTER 1

GENERAL NOTATION

For general information about operator algebras and for proofs of assertions made in this introductory section the books of S. Sakai [Sak97] and M. Takesaki [Tak03a] are referenced. In agreement with the usual convention we denote the natural, integer, real and complex numbers by $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$, respectively. For the set of natural numbers including zero we use $\mathbb{N}_0$ and for the subset of positive real numbers and the zero element $\mathbb{R}^+$. All Hilbert spaces are considered over the field $\mathbb{C}$ and symbolised by $\mathcal{H}$ (or $\mathcal{K}$). Their scalar products $\mathcal{H} \times \mathcal{H} \to \mathbb{C},\ (\xi, \eta) \mapsto \langle \xi, \eta \rangle$ are defined linear in the first and anti-linear in the second argument. The Banach space of all linear, bounded operators on a Hilbert space $\mathcal{H}$ with operator norm $||\cdot||$ is denoted by $\mathcal{B}(\mathcal{H})$. A *Banach algebra* is a Banach space $(E, ||\cdot||_E)$ together with an associative multiplication such that $||ab||_E \leq ||a||_E \cdot ||b||_E$. If additionally there is given an involution $E \to E,\ a \mapsto a^*$, we call E a *Banach* **-algebra.* A subalgebra $E \subseteq \mathcal{B}(\mathcal{H})$ is called an *operator algebra.* If $\dim \mathcal{H} = n < \infty$ then $\mathcal{B}(\mathcal{H})$ is isomorphic to a matrix algebra M_n. We make use of this fact freely and often write M_n instead of $\mathcal{B}(\mathcal{H})$.

1.1 A Few Facts about Operator Algebras

Topologies

Among the huge variety of topologies on Banach spaces there is only a small fraction that will be considered. First of all the norm topology induced by the norm of the Banach space. This topology is often referred to as *uniform topology*. The *weak topology* is given by the family of seminorms

$$E \ni x \mapsto |\varphi(x)|, \quad \varphi \in E^* \ .$$

If the Banach space has a unique predual E_*, the family of seminorms

$$E \ni x \mapsto |\varphi(x)|, \quad \varphi \in E_* \, ,$$

determines the so called *σ-weak topology*. If the Banach space is a Banach subspace of $\mathcal{B}(\mathcal{H})$, we define the *weak* and the *strong operator topology*, determined by the seminorms

$$A \ni x \mapsto |\langle x\xi, \eta\rangle|, \ \xi, \eta \in \mathcal{H} \quad \text{and} \quad A \ni x \mapsto |x\xi|, \ \xi \in \mathcal{H} \ .$$

To refer to limits with respect to one of the topologies mentioned above we write $|| \cdot ||$-, w-, $w*$-, *wop*- and *sop*-limit.

C^*-Algebras and von Neumann Algebras

A Banach*-algebra $\mathcal{A}$ is called a *C^*-algebra* if the condition $||xx^*|| = ||x|| \cdot ||x^*||$ is satisfied for all elements $x \in \mathcal{A}$. We call a C^*-algebra $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ a *von Neumann algebra* if $\mathcal{A}$ is *wop*-closed. If $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ is a C^*-algebra with identity, it is well known that a von Neumann algebra can be characterised by the condition $\mathcal{A} = \mathcal{A}''$ equivalently.
If the C^*-algebra $\mathcal{A}$ is the dual space of some Banach space $\mathcal{A}_*$, which we call the predual of $\mathcal{A}$, it is called a *W^*-algebra*. Since the predual of a C^*-algebra is unique

if it exists at all, speaking of 'the predual' of $\mathcal{A}$ is justified. Any W^*-algebra has a faithful representation as a von Neumann algebra $\pi(\mathcal{A})$ on a Hilbert space $\mathcal{H}$. Indeed, this last fact serves as an equivalent definition of a W^*-algebra. An element a in a C^*-algebra $\mathcal{A}$ is called *positive* if there exists an element $b \in \mathcal{A}$, such that $a = b^*b$. The set of positive elements is a convex cone and denoted by $\mathcal{A}_+$. As every finite dimensional Banach algebra has a unique predual automatically, we do not distinguish between a finite dimensional C^*- and W^*-algebra, and call both *finite algebra*.

Positive Functionals and States

A functional φ on a C^*-algebra $\mathcal{A}$ is called *positive* if $\varphi(a) \geq 0$ for every positive element $a \in \mathcal{A}_+$. If $\mathcal{A}$ is even a W^*-algebra, the local convex topology defined by the family of seminorms

$$\mathcal{A} \ni x \mapsto \varphi(x^*x)^{\frac{1}{2}}, \quad \varphi \in (\mathcal{A}_*)_+$$

is called the σ-strong topology . A positive functional φ that has norm $||\varphi|| = 1$ is a *state*. As to states on unital C^*-algebras even $\varphi(\mathbb{1}) = ||\varphi|| = 1$ is fulfilled. If a state is even strictly positive, that is $\varphi(a) > 0$ for every $0 \neq a \in \mathcal{A}_+$, we call it *faithful*. The set of states is a compact convex set with respect to the σ-weak topology on the C^*-algebra $\mathcal{A}^*$; we denote it by $\mathcal{S}(\mathcal{A})$. The extremal points of $\mathcal{S}(\mathcal{A})$ are referred to as *pure states*. Pure states on the C^*-algebra $\mathcal{B}(\mathcal{H})$ may be represented uniquely by Hilbert space vectors via $\mathcal{A} \ni a \mapsto \langle a\xi, \xi\rangle$ for $\xi \in \mathcal{H}$ with $||\xi|| = 1$. A positive linear functional on a W^*-algebra $\mathcal{A}$ that satisfies $\varphi(\sup a_i) = \sup \varphi(a_i)$ for every uniformly bounded directed set $\{a_i\}_{i \in I}$ of positive elements $a_i \in \mathcal{A}$ is said to be *normal*. The normal states on a W^*-algebra are exactly those which are σ-weakly continuous, or considering $\mathcal{A}_*$ as an linear subspace of $\mathcal{A}^*$, all states which belong to $\mathcal{A}_*$. A normal, faithful state $\varphi \in \mathcal{S}(\mathcal{A})$ on a W^*-algebra defines an inner product $\langle y, x\rangle_\varphi := \varphi(y^*x)$ and a norm $||x||_\varphi := (\langle x, x\rangle_\varphi)^{\frac{1}{2}} = \varphi(x^*x)^{\frac{1}{2}}$ for $x, y \in \mathcal{A}$.

1.2 Non-commutative Probability Theory

1.2.1 Non-commutative Probability Space

In classical probability theory a probability space (Ω, Σ, μ) consisting of a set Ω, a σ-algebra Σ and a probability measure μ serves to describe the random occurrences of events. Events correspond to subsets of the space Ω and those subsets contained in the σ-algebra Σ are exactly those events, which may occur in the random incidence we are observing. The measure μ quantifies the chance of a certain event to occur. By considering the commutative W^*-algebra $L^\infty(\Omega, \Sigma, \mu)$ of essentially bounded functions on Ω together with the state $\varphi_\mu(\cdot) = \int \cdot \, d\mu$ we are still able to describe the same events as before. Now, instead of subsets $A \in \Sigma$ the corresponding projections $\chi_A \in L^\infty(\Omega, \Sigma, \mu)$ represent events. On the other hand, every commutative W^*-algebra with a faithful state φ can be represented as a space $L^\infty(\Omega, \Sigma, \mu)$. This agreement between probability spaces (Ω, Σ, μ) and commutative W^*-algebras with normal state φ gives us reason to call the pair $(L^\infty(\Omega, \Sigma, \mu), \varphi_\mu)$ a (non)-commutative probability space. Then naturally, the next step in generalising probability spaces is allowing the W^*-algebra to be non-commutative.

Definition 1.1 Let $\mathcal{A}$ be a W^*-algebra and $\varphi \in \mathcal{S}(\mathcal{A})$ a normal, faithful state on $\mathcal{A}$. A pair $(\mathcal{A}, \varphi)$ is called a *non-commutative probability space.*

1.2.2 Product Spaces

Products of probability spaces are essential for constructing appropriate probability spaces for iterated stochastic experiments. Non-commutative product spaces are constructed by means of the tensor product of W^*-algebras. For finite products of probability spaces $(\mathcal{A}_i, \varphi_i)$, $i = 1, \ldots, n$, we consider the GNS representations $(\pi_\varphi, \mathcal{H}_{\varphi_i}, \xi_{\varphi_i})$. On the Hilbert space tensor product $\otimes_{i=1}^n \mathcal{H}_\varphi$ a von Neumann algebra $\bigotimes_{i=1}^n \mathcal{A}_i$ is generated by the elementary tensors $\pi_{\varphi_i}(a_1) \otimes \cdots \otimes \pi_{\varphi_n}(a_n)$. Together with the normal state

$$\otimes_{i=1}^n \varphi \colon \bigotimes_{i=1}^n \mathcal{A}_i \to \mathbb{C}, \; x_1 \otimes \cdots \otimes x_n \mapsto \varphi_1(x_1) \cdots \varphi_n(x_n)$$

we obtain the non-commutative probability space $(\otimes_{i=1}^n \mathcal{A}_i, \otimes_{i=1}^n \varphi)$. We refer to this state as the n-fold (tensor) product state $\otimes_{i=1}^n \varphi$ of the states $\varphi_i,\ i = 1, \ldots, n$. Each representation of the W^*-algebra $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ gained from faithful representations of the factors is a faithful representation of $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ [Tak03a]. Therefore, this constructive definition of the finite W^*-algebraic tensor product is unambiguous. If we have given a sequence of probability spaces $(\mathcal{A}_i, \varphi_i)_{i \in I}$, we regard the C^*-algebraic infinite tensor product $\otimes_{i \in I} \mathcal{A}_i$, where we consider $(\mathcal{A}_i)_{i \in I}$ as a sequence of C^*-algebras. We only regard countable sequences, either one-sided for $I = \mathbb{N}_0$ or two-sided for $I = \mathbb{Z}$. On $\otimes_{i \in I} A_i$ we define a infinite (tensor) product state $\otimes_I \varphi$ by

$$\otimes_{\mathbb{N}_0} \varphi(x_{i_0} \otimes \cdots \otimes x_{i_n} \otimes \mathbb{1}) = \varphi_{i_1}(x_{i_1}) \cdots \varphi_{i_n}(x_{i_n})$$

for all $n \in \mathbb{N}_0$ or for two-sided sequences by

$$\otimes_{\mathbb{Z}} \varphi(\mathbb{1} \otimes x_{i_{-n}} \otimes \cdots \otimes x_0 \otimes \cdots \otimes x_{i_n} \otimes \mathbb{1}) = \varphi_{i_{-n}}(x_{-i_n}) \cdots \varphi_0(x_0) \cdots \varphi_{i_n}(x_{i_n})$$

for all $n \in \mathbb{N}_0$. With respect to the state $\otimes\varphi$ we construct the GNS representation $(\pi_{\otimes\varphi_I}, \mathcal{H}_{\otimes\varphi_I}, \xi_{\otimes\varphi_I})$ of $\otimes_{i \in I} \mathcal{A}_i$. The bicommutant $\pi(\otimes_{i \in I} \mathcal{A}_i)''$ is a von Neumann algebra; it is called the infinite von Neumann tensor product. This construction depends on the sequence of states $(\varphi_i)_{i \in I}$. The pair $(\otimes_{i \in I} \mathcal{A}_i, \otimes\varphi)$ is called the infinite tensor product of the probability spaces $(A_i, \varphi_i)_{i \in I}$.

Example 1.2 Let $(L^\infty(\Omega_i, \Sigma_i, \mu_i), \varphi_{\mu_i}),\ \ i = 1, \ldots, n$ be a family of probability spaces. The finite tensor product of the W^*-algebras $L^\infty(\Omega_i, \Sigma_i, \mu_i)$ is given by

$$\bigotimes_{i=1}^n L^\infty(\Omega_i, \Sigma_i, \mu_i) \cong L^\infty\big(\prod_{i=1}^n \Omega_i, \otimes_{i=1}^n \Sigma_i, \otimes\mu_i\big).$$

Together with the product state $\otimes\varphi_{\mu_i}$ we obtain the finite product probability space $(\bigotimes_{i=1}^n L^\infty(\Omega_i, \Sigma_i, \mu_i), \otimes\varphi_{\mu_i})$.

1.2.3 Morphisms

Measurable transformations between probability spaces are among the main items under observation in classical probability theory. In particular, invertible transformations which preserve the involved measures proved to be worth studying. Given an invertible measure preserving transformation $\gamma\colon (\Omega,\Sigma,\mu)\to(\Omega,\Sigma,\mu)$ on a probability space (Ω,Σ,μ), we obtain from it a *-automorphism $T_\gamma\colon L^\infty(\Omega,\Sigma,\mu)\to L^\infty(\Omega,\Sigma,\mu),\ f\mapsto f\circ\gamma$ on the corresponding non-commutative probability space. The semigroups $(T_\gamma^n)_{n\in\mathbb{N}_0}$ arising from it are examples of so called dynamical systems $(L^\infty(\Omega,\Sigma,\mu),T_\gamma^n,\varphi_\mu)$, whose analysis has provided a substantial theory. However, the mere analysis of automorphisms would be too restricting. Already Markov processes elude here. We widen our scope by agreeing on the following definition.

Definition 1.3 A *morphism* $T\colon(\mathcal{A},\varphi)\to(\mathcal{B},\psi)$ of probability spaces $(\mathcal{A},\varphi)$ and $(\mathcal{B},\psi)$ is a completely positive operator $T\colon\mathcal{A}\to\mathcal{B}$, such that $T(\mathbb{1}_\mathcal{A})=\mathbb{1}_\mathcal{B}$ and $\psi(T(x))=\varphi(x)$ for all $x\in\mathcal{A}$.

If T is a *-isomorphism, then $T\colon(\mathcal{A},\varphi)\to(\mathcal{B},\psi)$ is called an *isomorphism*. Indeed, every *-isomorphism is completely positive and unital. On the other hand, if $T\colon(\mathcal{A},\varphi)\to(\mathcal{B},\psi)$ is an invertible morphism, we find

$$x^*x=T^{-1}T(x^*x)\geq T^{-1}(T(x^*)T(x))\geq T^{-1}T(x^*)T^{-1}T(x)=x^*x\ ,$$

where we made use of the Kadison-Schwartz inequality for the completely positive operator T. From the second equality

$$T^{-1}T(x^*x)=T^{-1}(T(x^*)T(x))$$

it follows that $T(x^*x)=T(x^*)T(x)$ for all $x\in\mathcal{A}$. To see the multiplicativity for arbitrary elements $x,y\in\mathcal{A}$ we define sesquilinear forms $s_\theta(x,y):=\theta(T(y^*x)-T(y)^*T(x))$ for all $\theta\in\mathcal{A}_*$. Applying the Cauchy-Schwartz inequality yields $T(y^*x)=T(y)^*T(x)$ for all $x,y\in\mathcal{A}$.

Accordingly a *-isomorphism $T\colon(\mathcal{A},\varphi)\to(\mathcal{A},\varphi)$ is called an *automorphism*. It is understood that $\varphi\circ T=\varphi$. In the beginning of this section we mentioned that every invertible measure preserving transformation on a probability space yields a *-automorphism. Conversely, every *-automorphism T on a probability space $(\mathcal{A},\varphi)$ with $\mathcal{A}$ a commutative W^*-algebra gives an invertible measure preserving transformation on some probability space (Ω, Σ, μ), where we identified $\mathcal{A}$ with the corresponding W^*-algebra of functions. Note that the above transformation is only unique μ-almost everywhere. We like to consider a morphism T on a probability space $(\mathcal{A}, \varphi)$ as a (discrete) *dynamical system* $(\mathcal{A}, T, \varphi)$ having in mind the discrete semigroup $(T^n)_{n\in\mathbb{N}_0}$. If T is an automorphism, we call $(\mathcal{A}, T, \varphi)$ *reversible dynamical system.*

Morphisms have two nice properties, which we would like to mention.

Property 1: Firstly, a morphism $T\colon (\mathcal{A}, \varphi) \to (\mathcal{B}, \psi)$ is faithful, meaning if $T(x^*x) = 0$ for an element $x \in \mathcal{A}$, then $x = 0$. T is inheriting this property from the faithful state φ: If $T(x^*x) = 0$ for an element $x \in \mathcal{A}$, then $\psi(T(x^*x)) = \varphi(x^*x) = 0$.

Property 2: Secondly, morphisms are normal, or equivalently, posses a preadjoint. For all $\theta \in \mathcal{A}_*$ with $0 \le \theta \le \psi$ we find $0 \le \theta \circ T \le \psi \circ T = \varphi$. Since φ is faithful, $\theta \circ T \in \mathcal{A}_*$. A net $(x_i^* x_i)_{i\in I}$ is converging σ-weakly to 0 if and only if $(x_i)_{i\in I}$ is converging σ-strongly to 0. Thus, by

$$|\theta \circ T(x)| \le \theta \circ T(x^*x) \le \varphi(x^*x)$$

it follows that $\theta \circ T$ is a σ-strongly continuous functional and as such σ-weakly continuous. Therefore $M_\psi := LH\{\theta \in \mathcal{A}_*\colon 0 \le \theta \le \psi\}$ is mapped into the predual $\mathcal{A}_*$ as well. By the Bipolar Theorem we learn that M_ψ is nothing less than a weakly and hence norm dense subset of $\mathcal{A}_*$.

To every morphism $T\colon (\mathcal{A}, \varphi) \to (\mathcal{B}, \psi)$ we can associate an operator $\overline{T}\colon \mathcal{H}_\varphi \to \mathcal{H}_\psi$ on the Hilbert spaces $\mathcal{H}_\varphi$ and $\mathcal{H}_\psi$ of the GNS representation of $\mathcal{A}$ and $\mathcal{B}$, respectively. $\overline{T}$ is given as extension of the operator $\overline{T}\colon \mathcal{A}\mathcal{H}_\varphi \to \mathcal{B}\mathcal{H}_\psi,\ a\mathbb{1}_\varphi \mapsto T(a)\mathbb{1}_\psi$ defined on the dense subspaces $\mathcal{A}\mathcal{H}_\varphi$ and $\mathcal{B}\mathcal{H}_\psi$. Restricted to those $\overline{T}$ is a contraction

$$||\overline{T}(a\mathbb{1}_\psi)||_\psi = \psi(T(a)^*T(a)) \le \psi(T(a^*a)) = \varphi(a^*a) = ||a\mathbb{1}_\varphi||_\varphi\ ,$$

where we have used the Kadison-Schwartz inequality. Therefore we can extend $\overline{T}\colon \mathcal{A}\mathcal{H}_\varphi \to \mathcal{B}\mathcal{H}_\psi$ on the closures. It was shown by B. Kümmerer [Küm] that for morphisms $T\colon (\mathcal{A}, \varphi) \to (\mathcal{A}, \varphi)$ the adjoint $\overline{T}^*$ comes from an underlying morphism

$T^+ : (\mathcal{A}, \varphi) \to (\mathcal{A}, \varphi)$ if and only if T commutes with the modular automorphism group $(\sigma_t^\varphi)_{t \in \mathbb{R}}$. If that is the case, the operator T^+ is called φ-adjoint and is uniquely determined by $\varphi(xT(y)) = \varphi(T^+(x)y)$ for all $x, y \in \mathcal{A}$. If T is already an automorphism, then $T^+ = T^{-1}$.

1.2.4 Conditional Expectation

Conditional expectations are an important concept of classical probability theory. Depending on the perspective they either extend a measure μ on a sub-σ-algebra $\Sigma_0 \subseteq \Sigma$ to a measure on the larger σ-algebra Σ or project Σ-measurable functions on to Σ_0-measurable functions. The crucial point in proving the existence of conditional expectations in classical probability spaces is the Radon-Nikodym Theorem. Although there are Radon-Nikodym Theorems for states on W^*-algebras, they do not meet our requirements. However, a projection P on a C^*-subalgebra $\mathcal{A}_0 \subseteq \mathcal{A}$ of norm one is equipped with the following properties for all $a, b \in \mathcal{A}_0$ and $x \in \mathcal{A}$ automatically:

1. $P(x^*x) \geq 0$,
2. $P(axb) = aP(x)b$,
3. $P(x^*)P(x) \leq P(x^*x)$.

Property 2) is referred to as module property and the reason for the following definition. Note that P is a completely positive map.

Definition 1.4 Let $(\mathcal{A}_0, \varphi)$ and $(\mathcal{A}, \psi)$ be non-commutative probability spaces. A projection of norm one $P : \mathcal{A} \to \mathcal{A}_0$, such that $\varphi \circ P = \psi$, is called *conditional expectation* onto $\mathcal{A}_0$.

On the GNS Hilbert space $\mathcal{H}_\varphi$ a conditional expectation induces an orthogonal projection P_φ by defining it on the dense subspaces $\mathcal{A}\mathcal{H}_\psi$ and $\mathcal{A}_0\mathcal{H}_\varphi$ as $P_\varphi : \mathcal{A}\mathcal{H}_\psi \to \mathcal{A}_0\mathcal{H}_\varphi$, $a\mathbb{1}_\psi \mapsto P(a)\mathbb{1}_\varphi$. From this we conclude that P is uniquely determined by the condition $\varphi \circ P = \varphi$, which justifies speaking of 'the' conditional expectation. Being a unital, completely positive operator, the conditional expectation $P : \mathcal{A} \to \mathcal{A}_0$ is a

morphism from $(\mathcal{A}, \psi)$ to $(\mathcal{A}_0, \varphi)$, and as such a normal operator.
The only drawback we have to suffer is, not every W^*-subalgebra possesses a conditional expectation. However, in [Tak03b] M. Takesaki proved a theorem which states that a conditional expectation on a W^*-subalgebra $\mathcal{A}_0 \subseteq \mathcal{A}$ exists if and only if $\mathcal{A}_0$ is left globally invariant by the modular automorphism group $(\sigma_t^{\varphi})_{t\in\mathbb{R}}$.

1.2.5 Random Variables

Physicists usually approach effects of random events in a system by considering the system to be embedded into a larger, surrounding system, which influences the smaller one. The former and the latter system are represented by probability spaces (Ω, Σ, μ) and (Ω', Σ', μ'), respectively, and the influence by a random variable $X\colon (\Omega', \Sigma', \mu') \to (\Omega, \Sigma, \mu)$, such that μ is the image measure $\mu = \mu' \circ X^{-1}$. The influence X is perceived or detected via the so called observable, a random variable $f\colon (\Omega, \Sigma, \mu) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On the non-commutative probability spaces $L^\infty(\Omega, \Sigma, \mu)$ and $L^\infty(\Omega', \Sigma', \mu')$ the random variable induces an injective *-homomorphism $i(X)\colon L^\infty(\Omega, \Sigma, \mu) \to L^\infty(\Omega', \Sigma', \mu'),\ f \mapsto f \circ X$, such that $\varphi_{\mu'} \circ i(X) = \varphi_\mu$. The algebra $L^\infty(\Omega, \Sigma, \mu)$ is often referred to as algebra of observables. From that perspective the random variable X maps observables of the smaller system to observables of the larger. The theory is rounded up by the fact that, indeed, every injective *-homomorphism between $W*$-algebras of functions on standard Borel spaces is induced by a random variable [Acc76].

Definition 1.5 Let $(\mathcal{A}, \varphi)$ be an non-commutative probability space, $\mathcal{A}_0$ a W^*-algebra and $J\colon \mathcal{A}_0 \to \mathcal{A}$ an injective *-homomorphism, such that the conditional expectation $P\colon \mathcal{A} \to \mathcal{A}_0$ exists. Then J is called *random variable*.

1.2.6 Stochastic Processes

The next necessary concept is that of a stochastic process. On a first glance being nothing else than a family of random variables $X_t : \Omega \to \mathbb{R}\,,\ t \in \mathbb{T}$ on a probability space, generalisation poses no difficulty.

Definition 1.6 Let $(\mathcal{A}, \varphi)$ be a probability space and $i_t \colon \mathcal{A}_0 \to (\mathcal{A}, \varphi)\,,\ t \in \mathbb{T}$, a family of random variables with $\mathbb{T} = \mathbb{N}_0, \mathbb{Z}$ or $\mathbb{R}$. We call $(\mathcal{A}, \varphi, (i_t)_{t\in\mathbb{T}}; A_0)$ a *stochastic process*.

Serving as models for a huge variety of applications, the class of so called stationary processes is of importance in classical probability theory. The intuitive idea of stationarity is that the random mechanism underlying the process is not changing in time. Mathematically this means that all finite marginal distributions are equal.

Definition 1.7 A stochastic process $(\mathcal{A}, \varphi, (i_t)_{t\in\mathbb{T}}; A_0)$ is called *stationary* if

$$\varphi(i_{t_1}(x_1)\cdots i_{t_n}(x_n)) = \varphi(i_{t_1+s}(x_1)\cdots i_{t_n+s}(x_n))$$

with $n \in \mathbb{N}_0$, $x_1, \ldots, x_n \in \mathcal{A}$, $t_1, \ldots, t_n \in \mathbb{T}$ and $s > 0$ arbitrarily. In particular $\varphi \circ i_0 = \varphi \circ i_n$ for all $n \in \mathbb{N}_0$.

Stochastic processes $X_t \colon \Omega \to \mathbb{R}$ may be also regarded as families of random functions. For every $\omega \in \Omega$ we can consider the function $\overline{X} \colon \mathbb{T} \to \mathbb{R}$, $\overline{X}(t) := X_t(\omega)$, the so called path or trajectory. The space of trajectories $\mathbb{R}^{\mathbb{T}}$ becomes a probability space by the Kolmogorov construction and the process $\chi_t \colon \mathbb{R}^{\mathbb{T}} \to \mathbb{R}$, $\omega \mapsto \omega_t$ of coordinate functions has the same joint distribution as the original process. The time shift $\chi_t = \chi_0 \circ \sigma_t,\ t \in \mathbb{T}$, induces a time translation on $\mathbb{R}^{\mathbb{T}}$.

Definition 1.8 A stochastic process $i_t \colon \mathcal{A}_0 \to (\mathcal{A}, \varphi)\,,\ t \in \mathbb{T}$, is represented by a *time translation* if there is given a family of *-homomorphisms $\hat{T}_t \colon (\mathcal{A}, \varphi) \to (\mathcal{A}, \varphi)$, such that for all $s, t \in \mathbb{T}$:

i) $\hat{T}_0 = \mathbb{1}$ and $\hat{T}_{s+t} = \hat{T}_s \hat{T}_t$,

ii) $i_{t+s} = \hat{T}_t i_s$.

Time translations exist for many processes [AFL82], in particular, if they are stationary.

If $i_t \colon \mathcal{A}_0 \to (\mathcal{A}, \varphi)$ is a stochastic process, we denote by $\mathcal{A}_I$ the algebra generated by $\{\, i_t(a) \colon a \in \mathcal{A}_0, t \in I \,\}$. We call a process *minimal* if $\mathcal{A} = \mathcal{A}_{\mathbb{T}}$.

Assertion 1.2.1 *The conditional expectation on $\mathcal{A}_I$ exists.*

Proof. For every $t \in \mathbb{T}$ the conditional expectation P_t on $\mathcal{A}_{\{t\}}$ exists. Therefore i_t commutes with the modular automorphism group $(\sigma_t^{\varphi})_{t\in\mathbb{R}}$ and so does the set $\{\, i_t \colon t \in I \,\}$ and the algebra $\mathcal{A}_I$ generated by it. □

1.2.7 Markov Processes

Markov processes are the simplest stochastic processes with dependent random variables. They are often described graphically as processes which do not posses a memory. This idea is aimed at the fact that for a Markov process events in the future only depend on the outcome in the present. Classically, the information about the events of a process $X_t \colon (\Omega, \Sigma, \mu) \to (\Omega_0, \Sigma_0)$ at a time t_0 is expressed by the σ-algebra $\Sigma_{\{t_0\}}$ generated by the random variable X_{t_0}. In the same manner the σ-algebra generated by $\{X_t \colon t \leq t_0\}$ describes the information from the past up to the time t_0 and is denoted by $\Sigma_{t_0]}$. Of course, to formulate dependence we need the conditional expectations $P_{\{t_0\}}$ and $P_{t_0]}$, which go with them. Then the dependence of the Markov process is just given by $P_{t_0]}(X_t) = P_{\{t_0\}}(X_t)$ for all $t \geq t_0$. For a non-commutative stochastic process we define the Markov property in corresponding terms analogously.

Definition 1.9 A stochastic process $i_t \colon \mathcal{A}_0 \to (\mathcal{A}, \varphi)$ is called a *Markov process* if

$$P_{t]}(x) = P_t(x) \quad \text{for all } x \in \mathcal{A}_{[t} \ .$$

1.2.8 Transition Operators

Dealing with dependent random variables, especially those of Markov processes, produces the notion of transition matrices and operators naturally. The entries of a transition matrix are the probabilities of the occurrence of a transition of one state to another. Between any pair of random variables $i \colon \mathcal{A}_0 \to (\mathcal{A}, \varphi)$ and $j \colon \mathcal{A}_0 \to (\mathcal{A}, \varphi)$ we can define an operator ${}_iT_j \colon \mathcal{A}_0 \to \mathcal{A}_0,\ a \mapsto i^{-1}P_i j(a)$, and think of it as the transition operator between the random variables i and j. As a composition of i^{-1}, P_i and j the operator ${}_iT_j$ is completely positive and unital. Additionally, ${}_iT_j$ maps the two states $\varphi_i := \varphi \circ i$ and $\varphi_j := \varphi \circ j$ induced by i and j onto each other:

$$\varphi_i({}_iT_jx) = \varphi \circ i(i^{-1}P_ij(x)) = \varphi(P_ij(x)) = \varphi(j(x)) = \varphi_j(x) \ .$$

Hence, ${}_iT_j\colon (\mathcal{A}_0, \varphi_j) \to (\mathcal{A}_0, \varphi_i)$ is a morphism from $(\mathcal{A}_0, \varphi_j)$ to $(\mathcal{A}_0, \varphi_i)$. If we are given a stochastic process $i_t\colon \mathcal{A}_0 \to (\mathcal{A}, \varphi)$, we define the transition operators accordingly.

Definition 1.10 Let $i_n\colon \mathcal{A}_0 \to (\mathcal{A}, \varphi)$, $n \in \mathbb{T}$, be a stochastic process. The operators defined by

$$T_n\colon \mathcal{A}_0 \to \mathcal{A}_0,\ a \mapsto i_0^{-1}P_0i_n(a)$$

are called *transition operators*.

The properties of the transition operator of two random variables passes on to those of a stochastic process; $\varphi_n \circ T_n = \varphi_0$ follows for the induced states $\varphi_k = \varphi \circ i_k,\ k \in \mathbb{T}$, too. If the stochastic process $i_t\colon \mathcal{A}_0 \to (\mathcal{A}, \varphi)$ is stationary, then

$$\varphi_0 \circ T_n(x) = \varphi_n(x) = \varphi_0 \circ i_0(x) = \varphi \circ i_0(x) = \varphi_0(x) \ .$$

Thus, φ_0 is an invariant state of the transition operators T_n. In classical probability theory there is a correspondence between semigroups (of probability kernels) and Markov processes. As for non-commutative Markov processes, we do not have a correspondence in general. If $i_n\colon \mathcal{A}_0 \to (\mathcal{A}, \varphi)$, $n \in \mathbb{N}_0$, is a Markov process the transition operators T_n may be written as products of transition operators of two random variables at successive times

$$\begin{aligned} T_n = i_0^{-1}P_0i_n &= i_0^{-1}P_0P_{1]}\cdots P_{n]}i_n \\ &= i_0^{-1}P_0i_1i_1^{-1}P_1\cdots P_ni_n = {}_{i_0}T_{i_1}\,{}_{i_1}T_{i_2}\cdots{}_{i_{n-1}}T_{i_n} \ . \end{aligned}$$

We call the Markov process homogeneous if ${}_{i_n}T_{i_{n+1}} = T_1$ for all $n \in \mathbb{N}_0$. Obviously, on that condition $(T_n)_{n\in\mathbb{N}_0}$ is a semigroup. Also, if $i_n\colon \mathcal{A}_0 \to (\mathcal{A}, \varphi)$ is a stochastic Markov process that possesses a time translation $(\hat{T}_n)_{n\in\mathbb{N}_0}$, such that $\varphi \circ \hat{T}_n = \varphi$

for all $n \in \mathbb{N}_0$, then $(T_n)_{n\in\mathbb{N}_0}$ is a semigroup:

$$\begin{aligned}T_s \cdot T_t &= i_0^{-1}P_0 i_s i_0^{-1} P_0 i_t = i_0^{-1}P_0\hat{T}_s i_0 i_0^{-1}P_0\hat{T}_t i_t = i_0^{-1}P_0\hat{T}_s P_0 \hat{T}_t i_t \\ &= i_0^{-1}P_0 P_s \hat{T}_s \hat{T}_t i_0 = i_0^{-1}P_0 P_{s]}\hat{T}_{s+t} i_0 = i_0^{-1}P_0\hat{T}_{s+t}i_0 = T_{s+t} \ .\end{aligned}$$

We have used the Markov property and the equation $P_s\hat{T}_s = \hat{T}_s P_0$, which follows from

$$\begin{aligned}\varphi(x\hat{T}_s P_0 y) = \varphi(P_s x \hat{T}_s P_0 y) &= \varphi((P_0\hat{T}_{-s}P_s x)(P_0 y)) \\ &= \varphi((P_s x)(P_s\hat{T}_s y)) = \varphi(xP_s\hat{T}_s y) \ .\end{aligned}$$

Definition 1.11 Let $(T_n)_{n\in\mathbb{N}_0}$ be a semigroup of normal, completely positive operators on a W^*-algebra. We call $(T_n)_{n\in\mathbb{N}_0}$ a *non-commutative Markov semigroup* if

1. $T_0 = id$,
2. $T_n(\mathbb{1}) = \mathbb{1}$ for all $n \in \mathbb{N}_0$.

A normal state $\varphi \in \mathcal{S}(\mathcal{A})$ is *invariant* if $\varphi \circ T_n = \varphi$ for all $n \in \mathbb{N}_0$.

1.2.9 One-sided Commutative Coupling

The situation we are going to describe in this section is important to us later, as we would like to understand the coupling from the past algorithm from the perspective, we have taken here. Let $(\mathcal{A}, \varphi)$ and $(\mathcal{C}, \psi)$ be finite dimensional commutative probability spaces which we think of as algebras of functions $L^\infty(A, \Sigma_A, \mu_\varphi)$ and $L^\infty(C, \Sigma_C, \mu_\psi)$. We will often denote those algebras by $L^\infty(A, \mu_\varphi)$ and $L^\infty(C, \mu_\psi)$, as for finite sets the σ-algebras are given by the power sets. Without loss of generality we assume that the measures μ_φ and μ_ψ are strictly positive, by considering their support in A and C, respectively. Further, let $J\colon (\mathcal{A}, \varphi) \to (\mathcal{A} \otimes \mathcal{C}, \varphi \otimes \psi)$ be a random variable. As morphisms on commutative probability spaces come from transformations on a probability space so do random variables. Namely, if $J\colon L^\infty(A, \Sigma_A, \mu_\varphi) \to L^\infty(A, \Sigma_A, \mu_\varphi) \otimes L^\infty(C, \Sigma_C, \mu_\psi)$ is a random variable, there exists a surjective measure preserving transformation $\gamma\colon A \times C \to A$, such that $J(f)(x) = f(\gamma(x, y))$ for all $f \in L^\infty(A, \Sigma_A, \mu_\varphi)$. Conversely, for every surjective, measure preserving transformation $\gamma\colon A \times C \to A$ we define a random variable $J\colon L^\infty(A, \Sigma_A, \mu_\varphi) \to L^\infty(A, \Sigma_A, \mu_\varphi) \otimes L^\infty(C, \Sigma_C, \mu_\psi)$ by $J(f) = f \circ \gamma$. Consider the transformation $\gamma\colon A \times C \to A$ as a process on A that is controlled by C. By the mapping γ we are given a set of mappings $\{\, \gamma(\cdot, c)\colon A \to A \,\}_{c \in C}$, that are chosen with probability $P(\gamma(\cdot, c)) = \mu_\psi(c)$. We illustrate this idea in Figure 1.1. We denote

$$\begin{tikzcd}[column sep=small]
A & \times & C & \times & C & \times & C & \times & \cdots \\
A & \times & C \arrow[ull, "\gamma"'] & \times & C \arrow[ull] & \times & C \arrow[ull] & \times & \cdots
\end{tikzcd}$$

Figure 1.1: One-sided coupling scheme

the product spaces $C^{\mathbb{N}_0}$ by C^+ and $A \times C^+$ by A^+. On the space A^+ we define a measure $\otimes\mu$ which is determined on cylinder sets

$$C_{a, c_1, \ldots, c_n} := \left\{\, \omega \in A^+ \colon \omega_0 = a, \omega_1 = c_1, \ldots, \omega_n = c_n \,\right\}$$

by $\otimes\mu(\omega) = \mu_\varphi(a)\mu_\psi(c_1) \cdots \mu_\psi(c_n)$. Then the transformation on A^+ is given by

$$\gamma^+ \colon A^+ \to A^+, \quad \begin{bmatrix} a & & \\ c_1 & c_2 & \ldots \end{bmatrix} \mapsto \begin{bmatrix} \gamma(a, c_1) & & \\ c_2 & c_3 & \ldots \end{bmatrix}.$$

1.2.10 One-sided Non-commutative Coupling

Let $(\mathcal{A}, \varphi)$ and $(\mathcal{C}, \psi)$ be non-commutative probability spaces and $J\colon (\mathcal{A}, \varphi) \to (\mathcal{A} \otimes \mathcal{C}, \varphi \otimes \psi)$ a random variable. As before we would like to interpret J as a coupling of a process on $\mathcal{A}$ to a process on $\mathcal{C}$. Only this times the random variable J is a mapping from $\mathcal{A}$ into the tensor product $\mathcal{A} \otimes \mathcal{C}$. We denote by $(\mathcal{C}^+, \psi^+)$ the tensor product of probability spaces $(\bigotimes_{n \in \mathbb{N}_0} \mathcal{C}, \otimes\psi)$. On the product space $\mathcal{A} \otimes \mathcal{C}^+$

$$\begin{tikzcd}
\mathcal{A} \arrow[d, "J"] & \otimes & \mathcal{C} \arrow[dr] & \otimes & \mathcal{C} \arrow[dr] & \otimes & \mathcal{C} \arrow[dr] & \otimes & \cdots \\
\overbrace{\mathcal{A} \otimes \mathcal{C}} & & \otimes & \mathcal{C} & \otimes & \mathcal{C} & \otimes & \mathcal{C} & \otimes \cdots
\end{tikzcd}$$

Figure 1.2: One-sided coupling scheme

we define a mapping $\hat{T}$ by

$$\hat{T}\colon \mathcal{A} \otimes \mathcal{C}^+ \to \mathcal{A} \otimes \mathcal{C}^+,\ a \otimes c \mapsto J(a) \otimes c \ .$$

The mapping $\hat{T}$ is an injective *-homomorphism. From it we construct a family of injective *-homomorphisms

$$i_n\colon \mathcal{A} \to \mathcal{A} \otimes \mathcal{C}^+,\ a \mapsto \hat{T}^n \circ i(a)\ ,$$

where $i\colon \mathcal{A} \to \mathcal{A} \otimes \mathcal{C}^+,\ a \mapsto a \otimes \mathbb{1}$ is the injection of $\mathcal{A}$ into $\mathcal{A} \otimes \mathcal{C}$. With J being a random variable, there exists a conditional expectation P on the algebra $J(\mathcal{A}) \subseteq \mathcal{C}^+$. We have already mentioned that this is equivalent to $J(\mathcal{A})$ being left invariant by the modular automorphism group $(\sigma_t^{\varphi\otimes\psi})_{t\in\mathbb{R}}$. From $(\sigma_t^{\varphi\otimes\psi})_{t\in\mathbb{R}}$ we obtain an automorphism group on $\mathcal{A}$ by $J^{-1}\sigma_t^{\varphi}J$. As this group satisfies the modular condition, the uniqueness of the modular automorphism group yields

$$\sigma_t^{\varphi} = J^{-1}\sigma_t^{\varphi}J \quad \text{or} \quad J\sigma_t^{\varphi} = \sigma_t^{\varphi}J\ .$$

Commuting with J the modular automorphism group $(\sigma_t^{\varphi\otimes\psi^+})_{t\in\mathbb{R}}$ commutes with the operator $\hat{T}$ and its powers as well. Therefore, the conditional expectation $P_{n]}$ exists, and the family of injective *-homomorphisms $i_n\colon (\mathcal{A},\varphi) \to (\mathcal{A}\otimes\mathcal{C}^+, \varphi\otimes\psi^+)$ is a stochastic process.

Observation 1 $i_n\colon \mathcal{A} \to \mathcal{A} \otimes \mathcal{C}^+$ *is a stationary Markov process.*

Again we can decompose it into a time shift on $\mathcal{C}^+$ - only this time it is the tensor right shift - and the random variable J.

1.2.11 Two-sided Commutative Coupling

Let $(\mathcal{A}, \varphi)$ and $(\mathcal{C}, \psi)$ be finite dimensional, commutative probability spaces again. As before we consider them as algebras of functions $L^\infty(A, \Sigma_A, \mu_\varphi)$ and $L^\infty(C, \Sigma_C, \mu_\psi)$ on finite sets A and C. Automorphisms on $(\mathcal{A}\otimes\mathcal{C}, \varphi\otimes\psi)$ originate in bijective measure preserving transformations $\gamma\colon A \times C \to A \times C$ on the probability spaces $(A \times C, \mu_\varphi \otimes \mu_\psi)$. If $T\colon (\mathcal{A}\otimes\mathcal{C}, \varphi\otimes\psi) \to (\mathcal{A}\otimes\mathcal{C}, \varphi\otimes\psi)$ is an automorphism, there exists a bijective measure preserving transformation $\gamma\colon A \times C \to A \times C$, such that $T \circ f = f \circ \gamma$ for all $f \in L^\infty(A \times C, \mu_\varphi \otimes \mu_\psi) \simeq (\mathcal{A} \otimes \mathcal{C}, \varphi \otimes \psi)$. It takes hardly by surprise that every bijective measure preserving transformation $\gamma\colon A \times C \to A \times C$ provides an automorphism

$$T_\gamma\colon (\mathcal{A} \otimes \mathcal{C}, \varphi \otimes \psi) \to (\mathcal{A} \otimes \mathcal{C}, \varphi \otimes \psi),\ f \otimes g \mapsto f \otimes g \circ \gamma\ .$$

Having the idea of coupling in mind, it is instructive to present a bijective measure preserving transformation $\gamma\colon A \times C \to A \times C$ as is done in Figure 1.3.

$$\begin{array}{c} \cdots \times C \times C \times C \times \left.\begin{matrix} A \\ \\ C \end{matrix}\right\}\gamma \times C \times C \times C \times \cdots \\ \xleftarrow[\sigma]{\qquad\qquad} \end{array}$$

Figure 1.3: Two-sided coupling scheme (version 1)

We denote the product space $C^{\mathbb{Z}}$ by $\hat{C}$ and by $\otimes\mu_\psi$ the product measure on $\hat{C}$. On $(\hat{C}, \otimes\mu_\psi)$ we can define a natural measure preserving transformation - the Bernoulli shift σ. The coupling depicted in Figure 1.3 is assembled from the Bernoulli shift σ and the transformation

$$\gamma\colon A \times \hat{C} \to A \times \hat{C},\ \begin{bmatrix} & & a & & \\ \cdots & c_{-1} & c_0 & c_1 & \cdots \end{bmatrix} \mapsto \begin{bmatrix} & & \gamma_A(a, c_0) & & \\ \cdots & c_{-1} & \gamma_C(a, c_0) & c_1 & \cdots \end{bmatrix}\ ,$$

where $A \times C \ni \gamma(a,c) = (\gamma_A(a,c), \gamma_C(a,c))$ with $\gamma_A(a,c) \in A$ and $\gamma_C(a,c) \in C$. The composition $\sigma \circ \gamma$ results in the transformation

$$\hat{\gamma} \colon A \times \hat{C} \to A \times \hat{C},$$
$$\begin{bmatrix} & & a & & \\ \cdots & c_{-1} & c_0 & c_1 & \cdots \end{bmatrix} \mapsto \begin{bmatrix} & & \gamma_A(a,c_0) & & \\ \cdots & \gamma_C(a,c_0) & c_1 & c_2 & \cdots \end{bmatrix}.$$

More than it is in the one-sided case, it is now justified to speak of a coupling. The basic process is the Bernoulli process on the product space $(\hat{C}, \otimes\mu_\psi)$ whose outcomes inflict a process on (A, μ_φ) by means of the local transformation γ. The one-sided case, however, is better considered as a controlled process on A, where the one-sided sequence of elements $c \in C$ are control parameters, which are forgotten after they have been processed. An alternative way of illustrating the transformation $\hat{\gamma}$ is given in Figure 1.4.

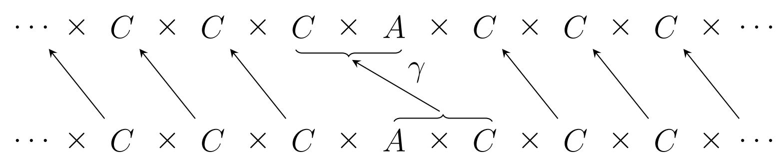


Figure 1.4: Two-sided coupling scheme (version 2)

1.2.12 Two-sided Non-commutative Coupling

Let $(\mathcal{A}, \varphi)$ and $(\mathcal{C}, \psi)$ be non-commutative probability spaces and $T \colon (\mathcal{A} \otimes \mathcal{C}, \varphi \otimes \psi) \to (\mathcal{A} \otimes \mathcal{C}, \varphi \otimes \psi)$ be an automorphism. Similarly to the cases we have examined before and inspired by their coupling schemes, we construct a coupling as in Figure 1.5. On the two-sided infinite tensor product $(\hat{\mathcal{C}}, \hat{\psi}) := (\otimes_{\mathbb{Z}} \mathcal{C}, \otimes\psi)$ of the probability space $(\mathcal{C}, \psi)$ we define the tensor right shift S. We extend S to the automorphism $\hat{S} := Id_{\mathcal{A}} \otimes S$ on $(A \otimes \hat{\mathcal{C}}, \varphi \otimes \hat{\psi})$ and the automorphism $T \colon (\mathcal{A} \otimes \mathcal{C}, \varphi \otimes \psi) \to (\mathcal{A} \otimes \mathcal{C}, \varphi \otimes \psi)$ trivially to an automorphism on $(\mathcal{A} \otimes \hat{\mathcal{C}}, \varphi \otimes \hat{\psi})$.

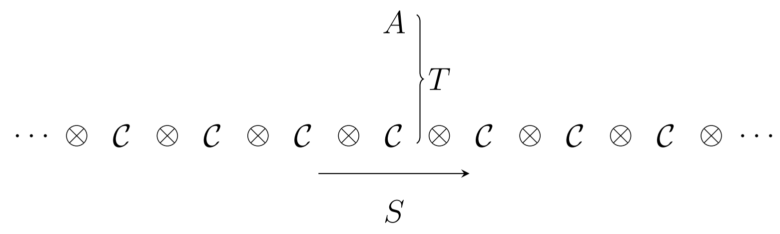


Figure 1.5: Two-sided coupling scheme (version 1)

Then we define the automorphism

$$\hat{T}\colon (\mathcal{A}\otimes\hat{\mathcal{C}}, \varphi\otimes\hat{\psi}) \to (\mathcal{A}\otimes\hat{\mathcal{C}}, \varphi\otimes\hat{\psi}),\ \hat{T} := T\circ\hat{S}\ .$$

As for automorphisms on commutative probability spaces we can illustrate the effect of the automorphism in a different way (Figure 1.6). The automorphism is a

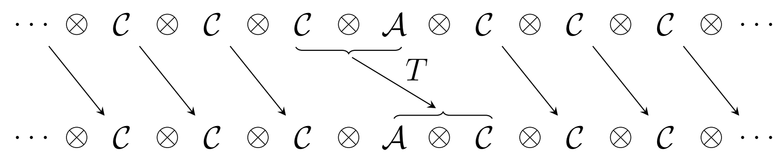


Figure 1.6: Two-sided coupling scheme (version 2)

time translation of the stochastic process defined by

$$i_n\colon (\mathcal{A},\varphi) \to (\mathcal{A}\otimes\hat{\mathcal{C}}, \varphi\otimes\hat{\psi}), a \mapsto \hat{T}^n \circ i(a)$$

where $i\colon (\mathcal{A},\varphi) \to (\mathcal{A}\otimes\hat{\mathcal{C}}, \varphi\otimes\hat{\psi}),\ a\mapsto a\otimes\mathbb{1}_{\hat{\mathcal{C}}}$ is the random variable that embeds the algebra $\mathcal{A}$ into the tensor product $\mathcal{A}\otimes\hat{\mathcal{C}}$.

Observation 2 *The family of injective *-homomorphisms*

$$i_n \colon (\mathcal{A}, \varphi) \to (\mathcal{A} \otimes \hat{\mathcal{C}}, \varphi \otimes \hat{\psi})$$

is a stationary Markov process.

Proof. The stationarity of the process follows from the fact that $\hat{T}$ and $\hat{S}$ leave the product state $\varphi \otimes \hat{\psi}$ invariant. For $t = 0$ the Markov property is verified by inspecting the algebra $\mathcal{A}_{[0}$ and the algebra $\mathcal{A}_{0]}$. Therefore $P_{0]}x = P_0 x$ for every $x \in \mathcal{A}_{[0}$. As $P_{t]} = \hat{T}^t P_{0]} \hat{T}^{-1}$, it follows

$$P_{t]} \hat{T}^t \hat{T}^s x = \hat{T}^t P_{0]} \hat{T}^s x = \hat{T}^t P_0 \hat{T}^s x = P_t \hat{T}^t \hat{T}^s x$$

for all $t \in \mathbb{Z}, s \in \mathbb{N}_0$ and $x \in \mathcal{A}_0$. □

1.2.13 Micromaser

The term micromaser denotes an experimental setup in quantum optics that has gathered some fame in the last twenty years [Vea00]. In this experiment excited atoms are brought to interact with photons. The former are so called Rydberg atoms - highly excited rubidium ^{85}Rb atoms. The photon field is a standing microwave in a resonator, which is realised as a superconducting cavity. This experiment serves as a fundamental prototype to investigate the interaction of photons and matter, as well as basic principles of quantum mechanics. Apart from that, after quantum optics came in the focal point of quantum information, the micro maser is employed to generate non-classical states like Fock states or - situated in the very heart of numerous questions - entangled states. A beam of Rydberg atoms is directed through a tiny opening in the cavity. By selecting only those atoms with a distinct velocity the interaction time t_0 between atom and electromagnetic field can be controlled. The frequency of the microwave is adjusted, such that a mode of the field is equal to the energy difference between excited and non excited state of the atoms. The flux of the beam is reduced sufficiently to allow only one atom passing through the cavity at a time. To simplify

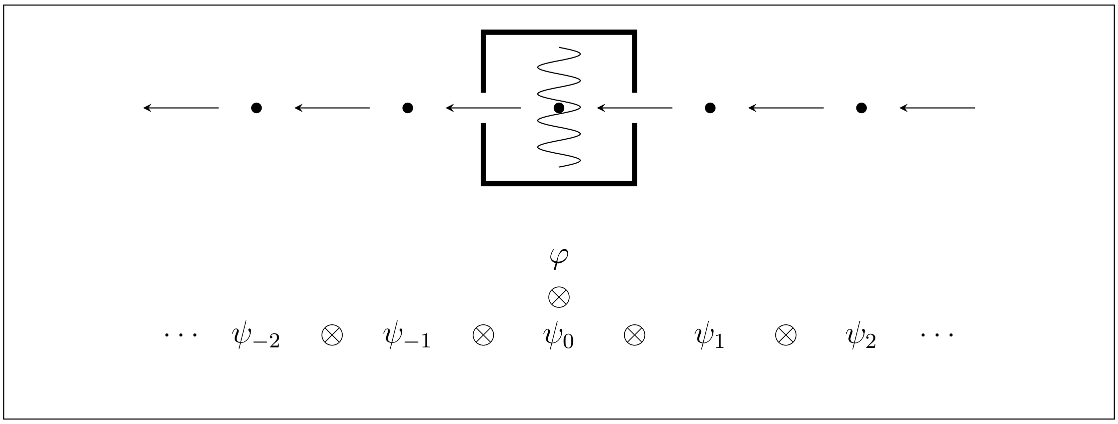
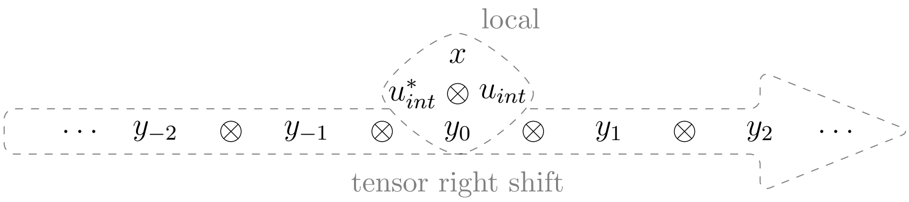


Figure 1.7: Micro maser

description we additionally assume that the distribution of the atoms along the beam is equidistant. Mathematically the mode of an electromagnetic field is a harmonic oscillator, whose algebra of observables is given by $\mathcal{A} = \mathcal{B}(\mathcal{H})$ either on the Hilbert space $\mathcal{H} = L^2(\mathbb{R})$ or $\mathcal{H} = \ell^2(\mathbb{N}_0)$, depending on the frame of description (position or energy representation). The algebra of observables of the beam of atoms, each of them considered as a two level system, is described by the infinite tensor product of matrix algebras M_2. Hence, the whole system is represented by the tensor product $\mathcal{A} \otimes \bigotimes_{\mathbb{Z}} M_2$, where the distinguished factor $\mathcal{A}$ is thought to be attached to the zeroth component of $\bigotimes_{\mathbb{Z}} M_2$. The dynamics on $\mathcal{A} \otimes \bigotimes_{\mathbb{Z}} M_2$ is given by an automorphism $\alpha \colon \mathcal{A} \otimes \bigotimes_{\mathbb{Z}} M_2 \to \mathcal{A} \otimes \bigotimes_{\mathbb{Z}} M_2$. Two parts comprise α: the local interaction between photon and atom in the cavity and the translation of the atoms.



The interaction between electromagnetic field and atom is given by a unitary automorphism

$$\alpha_{int}\colon \mathcal{A}\otimes M_2 \to \mathcal{A}\otimes M_2,\ x\otimes y \mapsto u_{int}^*(x\otimes y)u_{int}\ ,$$

where $u_{int} = e^{i\mathbb{H}t_0}$, $\mathbb{H}$ is the Hamiltonian and t_0 the interaction time. The movement of the atoms is described by the tensor shift S. The composition $\alpha := \alpha_{int}\circ S$ yields the full dynamics, i.e. the automorphism α. It is evident that the process

$$\mathcal{A} \ni a \mapsto \alpha^t \circ i(a) \in \mathcal{A}\otimes \bigotimes_{\mathbb{Z}} M_2\ ,\ t\in\mathbb{Z}\ ,$$

is a quantum Markov process, where $i\colon \mathcal{A}\to\mathcal{A}\otimes\bigotimes_{\mathbb{Z}} M_2$ is the injection $i(a) = a\otimes\mathbb{1}$. Instead of considering this process on the algebra of observables, a perspective called 'Heisenberg picture' in physics, we might as well describe it on the state space. This dual representation is usually referred to as 'Schrödinger picture'. We obtain it by adjoining all the involved automorphisms:

$$\alpha^*\colon \mathcal{S}(\mathcal{B}(\mathcal{H}))\otimes M_2 \to \mathcal{S}(\mathcal{B}(\mathcal{H}))\otimes M_2,\ \varphi\otimes\psi \mapsto (\varphi\otimes\psi)\circ\alpha = S^*\circ\alpha_{int}^*\circ(\varphi\otimes\hat{\psi})\,.$$

What we have ignored so far, is the precise description of the interaction by its Hamiltonian. As an approximation the so called Jaynes-Cummings model has proven to be adequate. Its Hamiltonian (in the energy representation) on $\ell^2(\mathbb{N}_0)\otimes\mathbb{C}^2$ is given by

$$\begin{aligned}
\mathbb{H} &= \hbar\omega_F a^*a\otimes\mathbb{1} + \mathbb{1}\otimes\frac{\hbar}{2}\omega_A\sigma_z + g\hbar(a+a^*)\otimes(\sigma_+ + \sigma_-)\\
&\approx \hbar\omega_F a^*a\otimes\mathbb{1} + \mathbb{1}\otimes\frac{\hbar}{2}\omega_A\sigma_z + g\hbar(a\otimes\sigma_+ + a^*\otimes\sigma_-)\\
&\approx \hbar\omega a^*a\otimes\mathbb{1} + \mathbb{1}\otimes\frac{\hbar}{2}\omega\sigma_z + g\hbar(a\otimes\sigma_+ + a^*\otimes\sigma_-)\ .
\end{aligned}$$

With ω_F being the frequency of the electromagnetic field, ω_A the frequency equivalent of the energy difference of the two atom levels and g the coupling constant, the Hamiltonian of the full system decomposes into the Hamiltonian of the field $\mathbb{H}_F = \hbar\omega_F a^*a\otimes\mathbb{1}$, the Hamiltonian of one atom $\mathbb{H}_A = \frac{\hbar}{2}\omega_A\sigma_z\otimes\mathbb{1}$ and the Hamiltonian of the interaction $\mathbb{H}_{int} = g\hbar(a + a^*) \otimes (\sigma_+ + \sigma_-)$. The second line results from an approximation called rotating wave approximation, usually applied if $|\omega_F - \omega_A|$ is much smaller than $\omega_F + \omega_A$. Assuming $\omega_A = \omega_F =: \omega$ yields the last line.

1.2.14 Scattering Theory

Whenever there is given a positive definite family of contractions $(T_n)_{n\in\mathbb{Z}}$, $T_0 = \mathbb{1}$, on a Hilbert space $\mathcal{H}_0$ there exists a unitary operator $\hat{T}$ on a Hilbert space $\mathcal{H}$, an isometry $i\colon \mathcal{H}_0 \to \mathcal{H}$ and an orthogonal projection $P_0\colon \mathcal{H} \to \mathcal{H}_0$, such that the following diagram commutes for all $n \in \mathbb{Z}$:

$$\begin{tikzcd} \mathcal{H}_0 \arrow[r, "T_n"] \arrow[d, "i"'] & \mathcal{H}_0 \\ \mathcal{H} \arrow[r, "\hat{T}^n"] & \mathcal{H} \arrow[u, "P_0"'] \end{tikzcd} .$$

The triple $(\mathcal{H}, \hat{T}, \mathcal{H}_0)$ is called a unitary dilation of the family $(T_n)_{n\in\mathbb{Z}}$. Conversely, if $(\mathcal{H}, (T_n)_{n\in\mathbb{Z}})$ has a unitary dilation, $(T_n)_{n\in\mathbb{Z}}$ is positive definite. Those results date back to the late 1960's ([LP67], [SNF70]) and serve as mathematical background in physics to describe an irreversible dynamic in a small system that emerges by reduction of a reversible process in a larger surrounding Hilbert space. If the family $(T_n)_{n\in\mathbb{Z}}$ is even a semigroup, i.e. $T_n = T_1^n$, the unitary dilation possesses a nice geometric structure. If $(\mathcal{H}, \hat{T}, \mathcal{H}_0)$ is a unitary dilation of such a semigroup $(T_n)_{n\in\mathbb{Z}}$, we denote the closed linear span $\left\{ \hat{T}^n\xi\colon \xi \in \mathcal{H}_0, n \in I \right\}$ for any subset $I \in \mathbb{Z}$ by $\mathcal{H}_I$ and the orthogonal projection from $\mathcal{H}$ onto $\mathcal{H}_I$ by $P_I\colon \mathcal{H} \to \mathcal{H}_0$.

Proposition 1.12 *Let $(\mathcal{H}, \hat{T}, \mathcal{H}_0)$ be a unitary dilation of a positive definite family of contractions $(T_n)_{n\in\mathbb{Z}}$. Then the following assertions are equivalent:*

a) $(T_n)_{n\in\mathbb{Z}}$ is a semigroup.

b) $\hat{T}^m P_0^\perp \hat{T}^n \xi$ is orthogonal to $\mathcal{H}_0$ for all $\xi \in \mathcal{H}_0$.

c) $P_{]-\infty,0]}(\xi) = P_0(\xi)$ for all $\xi \in \mathcal{H}_{[0,\infty[}$.

The geometric property b) allows us to construct a unitary dilation on the Hilbert space $\mathcal{H} := \mathcal{H}_0 \oplus \bigoplus_{\mathbb{Z}} \mathcal{H}_0$. Figure 1.8 captures the idea of the construction: where U_1 is an unitary on $\mathcal{H}_0 \oplus \mathcal{H}_0$ and S the right shift on $\oplus_{\mathbb{Z}}\mathcal{H}_0$. The unitary operator U_1 accrues from T by compensating for what is missing T to be an isometry and

$$\cdots \oplus \mathcal{H}_0 \oplus \left.\begin{matrix}\mathcal{H}_0\\ \oplus \\ \mathcal{H}_0\end{matrix}\right\} U_1 \oplus \mathcal{H}_0 \oplus \cdots$$

$$\xrightarrow[S]{\qquad\qquad}$$

Figure 1.8: Construction scheme of unitary dilation

modifying this isometry, such that it becomes unitary on $\mathcal{H}_0 \oplus \mathcal{H}_0$:

$$U_1 := \begin{pmatrix} T & -\sqrt{\mathbb{1} - TT^*} \\ \sqrt{\mathbb{1} - T^*T} & T^* \end{pmatrix} .$$

Explicitly the dilation $(\mathcal{H}, \hat{T}, \mathcal{H}_0)$ is given by $U := U_1 \circ S$, the insertion at the zeroth summand $i \colon \mathcal{H}_0 \to \mathcal{H},\ \xi \mapsto \xi \oplus 0$ and $P := i^*$ the projection onto the distinguished Hilbert space. If we consider the shift S as free evolution, U denotes the free evolution locally perturbed by U_1. From this point of view it is natural to ask if the dynamics becomes free again for the limit $n \to \infty$. We can try to answer this question by comparing the evolutions U and S by means of the so called wave operator: $\Phi_- := \lim_\infty S^{-n}U^n$.

Proposition 1.13 *The following assertions are equivalent.*

a) $\Phi_- := sop - \lim_{n\to\infty} S^{-n}U^n$ *exists and* $\Phi_-(\mathcal{H}) \subseteq \mathcal{H}_0^\perp$.

b) $sop - \lim_{n\to\infty} T^n = 0$.

c) There exists an isometry $\Phi_- \colon \mathcal{H}_0 \oplus \ell^2(\mathbb{Z}, \mathcal{H}_0) \to \ell^2(\mathbb{Z}, \mathcal{H}_0)$, *such that* $\Phi_-|_{\ell^2(\mathbb{N}_0, \mathcal{H}_0)} = Id$ *and* $S = \Phi_- U \Phi_-^{-1}$.

d) $\bigcup_{n\geq 0} U^{-n}(\ell^2(\mathbb{N}_0, \mathcal{H}_0))$ *is dense in* $\mathcal{H}$.

Motivated by the above theory of unitary dilations B. Kümmerer and H. Maassen [KM00] developed a scattering theory for Markov processes.

Scattering Theory for Markov Processes

Let $(\mathcal{A}, \varphi)$ and $(\mathcal{C}, \psi)$ be two probability spaces and $T\colon (\mathcal{A}\otimes\mathcal{C}, \varphi\otimes\psi) \to (\mathcal{A}\otimes\mathcal{C}, \varphi\otimes\psi)$ an automorphism. From it we construct a stationary Markov process as in section 1.2.12. Equivalence d) of Proposition 1.13 serves as a starting point and motivates the definition

$$\mathcal{A}_{out} := \overline{\bigcup \hat{T}^{-n}(\mathcal{C}_{[0,\infty)})}^{||\cdot||_\varphi} \subseteq \hat{A} \ .$$

All elements of $\hat{A}$ that will end in $\mathcal{C}_{[0,\infty)}$ eventually are gathered here. Note that on convex sets the topology induced by the norm $||\cdot||_\varphi$ is equivalent to the σ-weak topology. Let further Q be the conditional expectation onto the algebra $\mathcal{C}_{[0,\infty)}$.

Lemma 1.14 *The following assertions are equivalent for every $x \in \hat{A}$:*

a) $x \in \mathcal{A}_{out}$*;*

b) $\lim_{n\to\infty} ||Q \circ \hat{T}_n(x) - \hat{T}_n(x)||_\varphi = 0$*;*

c) $\lim_{n\to\infty} ||Q \circ \hat{T}_n(x)||_\varphi = ||x||_\varphi$*;*

d) The limit $||\cdot||_\varphi - \lim_{n\to\infty} S^{-n}\hat{T}_n(x)$ *exists and is an element of* $\mathcal{C}$*.*

Theorem 1.15 *The following assertions are equivalent:*

a) $\hat{A} = \mathcal{A}_{out}$*;*

b) $\lim_{n\to\infty} ||Q \circ \hat{T}_n(x)||_\varphi = ||x||_\varphi$ *for all* $x \in \hat{\mathcal{A}}$*;*

c) $\lim_{n\to\infty} ||Q \circ \hat{T}_n(x) - \hat{T}_n(x)||_\varphi = 0$ *for all* $x \in \hat{\mathcal{A}}$*;*

d) The limit $||\cdot||_\varphi - \lim_{n\to\infty} S^{-n}\hat{T}_n(x)$ *exists for all* $x \in \hat{A}$ *and is an element of* $\mathcal{C}$*;*

e) There exists an isomorphism $\Psi\colon (\hat{A}, \hat{\psi}) \to (\hat{C}, \hat{\psi})$ *with* $\Psi|_{\hat{C}_{[0,\infty)}} = Id$ *such that* $S\Psi = \Psi\hat{T}$*.*

Definition 1.16 Let $T\colon (\mathcal{A}\otimes\mathcal{C}, \varphi\otimes\psi) \to (\mathcal{A}\otimes\mathcal{C}, \varphi\otimes\psi)$ be an automorphism. We say T is asymptotically complete if T satisfies one of the equivalent conditions of the above theorem 1.15. Then the corresponding Markov process $(\mathcal{A}\otimes\mathcal{C}, \varphi\otimes\psi, i_n; \mathcal{A})$ where $i_n := \hat{T}^n \circ i$ is called asymptotically complete as well.

T. Lang [Lan03] has implemented this scattering theory to random variables $J\colon (\mathcal{A},\varphi) \to (\mathcal{A}\otimes\mathcal{C}, \varphi\otimes\psi)$ or, put differently, to one-sided processes. By identifying the one-sided process $(\mathcal{A}^+, \varphi\otimes\psi^+, T^+; \mathcal{A})$ where the injective *-homomorphism T^+ is given by

$$T^+\colon \mathcal{A}\otimes\mathcal{C}^+ \to \mathcal{A}\otimes\mathcal{C}^+,\ a\otimes c \mapsto J(a)\otimes c$$

with a two-sided process $(\hat{\mathcal{A}}, \varphi\otimes\hat{\psi}, \hat{T}, \mathcal{A})$ by extending T^+ to $\hat{T}$:

$$\hat{T}(c\otimes a) = c\otimes T^+(a)\ , \quad a\in\mathcal{A}^+,\ c\in\mathcal{C}^-,$$

all of the above assertions can be translated to one-sided processes equivalently. The analogous version of Theorem 1.15 reads as follows.

Theorem 1.17 *The following assertions are equivalent:*

a) $\hat{\mathcal{A}} = \mathcal{A}_{out}$*;*

b) $\lim_{n\infty} ||Q\circ\hat{T}_n(x)||_\varphi = ||x||_\varphi$ *for all* $x\in\hat{A}$*;*

c) $\lim_{n\infty} ||Q\circ\hat{T}_n(x) - \hat{T}_n(x)||_\varphi = 0$ *for all* $x\in\hat{A}$*;*

d) The limit $||\cdot||_\varphi - \lim_{n\to\infty} S^{-n}\hat{T}_n(x)$ *exists for all* $x\in\hat{A}$ *and is an element of* $\mathcal{C}$*;*

*e) There exists an injective *-homomorphism* $\Psi\colon (A^+,\varphi^+) \to (\hat{C},\hat{\psi})$ *with* $\Psi|_{C^+} = Id$ *and* $\Psi i^+(\mathcal{A}) \subseteq \hat{C}_{(-\infty,-1]}$*, such that* $S\circ\Psi = \Psi\circ J^+$*.*

Definition 1.18 We say the one-sided Markov process $(\mathcal{A}^+, \varphi\otimes\psi^+, T^+, i)$ is asymptotically complete if it conforms one of the equivalent conditions of the above theorem 1.17.

CHAPTER 2

GRAPHS AND STOCHASTIC MATRICES

Graphs are used to model and visualise many different problems and situations. Historically those reach from the famous Königsberg bridge problem to modern applications in the theory of automata and of course stochastic processes.

2.1 Graphs

A graph (V, E) consists of a set of vertices V and a set of edges E. Dealing with Markov processes so called directed graphs are common.

Definition 2.1 A *directed graph* is a graph, such that for every edge $e \in E$ there exists a starting point $s\colon E \to V$ and a target $t\colon E \to V$.

The number of incoming edges of a vertex $v \in V$ is called the indegree $In(v)$ and the number of outgoing edges the outdegree $Out(v)$. A graph (V, E) whose edges are coloured, i.e. there exists a set C and a mapping $c\colon E \to C$ which relates a colour $c(e) \in C$ to each edge $e \in E$, is called a coloured graph. If we can assign a mutual colour to an outgoing edge for all vertices simultaneously and uniquely, such that all edges are coloured, we call the graph road-coloured.

Definition 2.2 i) A directed graph (E, V) is called *c-graph* if there exist a set of colours and a mapping $c\colon E \to C$.

ii) A c-graph is called *road-coloured* if $\{\, c(e)\colon e \in s^{-1}(v) \,\} = C$ for all $v \in V$.

The colouring of a road-coloured graph is a systematic assignment of labels that allows giving global directions, i.e. directions independent of the vertices. The ordered colour $c \in C$ can be followed from every vertex $v \in V$ unambiguously. A special class of road-coloured graphs are those which posses so called synchronising words.

Definition 2.3 A *synchronising word* of a road-coloured graph (G, C) is a finite sequence $c_1 \cdots c_n$ of colours $c_i \in C$, $i = 1, \ldots, n$, such that for every sequence $e_1, \ldots, e_n$ of edges $e_i \in E$, $i = 1, \ldots, n$, with $t(e_i) = s(e_{i+1})$ for all $i = 1, \ldots, n-1$ and $c(e_i) = c_i$ for all $i = 1, \ldots, n$ the target vertex $t(e_n)$ is the same.

Ordering a synchronising word leads every path along the graph to one specific vertex. Hence, road-coloured graphs that posses synchronising words are interesting for the theory of automata, since synchronising words can be used to reset an automaton to a default state after the occurrence of an error.
Whether a graph, given he can be road-coloured at all, possesses a colouring with synchronising words, was an open problem for years. For obvious reasons it is referred to as the road-colouring problem. In 2007 A. N. Thratman [Tra09] proved it is possible indeed if the graph is aperiodic irreducible and has constant outdegree. We will return to this statement later on, as it has some relevance for our research as well. Modelling stochastic processes requires us to introduce weights to the edges. Those weights represent the frequency or the probability of the use of a certain edge on a path along the graph.

Definition 2.4 A *weighted* graph (G, ω) is a graph $G = (V, E)$ with a mapping $\omega\colon E \to [0, 1]$. A weighted graph is called *stochastic* if $\sum\limits_{s(e)=v} \omega(e) = 1$ for all vertices $v \in V$.

2.1.1 Synchronising Words and Asymptotic Completeness

As we have pointed out in Section 1.2.9 there is a one to one correspondence between surjective maps $\gamma\colon A \times C \to A$ on finite sets A, C and non-commutative random variables $J\colon L^\infty(A) \to L^\infty(A) \otimes L^\infty(C)$. With regard to this correspondence it is even possible to identify stricter and more precise equivalences. This research has been undertaken by T. Lang [Lan03] mainly. Some results can be found in [GKL06] as well. Before we proceed with stating those results, we list some easy and useful facts about graphs with synchronising words. We begin with a classification of the vertices of a graph, which not by chance bears a strong resemblance to the classification of transition matrices of Markov chains.

Definition 2.5 A directed graph $G = (V, E)$ is called *irreducible* if V is the smallest set $\emptyset \neq W$ of vertices in V, such that there exists no path that leaves W.

This property is referred to as strongly connected in literature as well.

Definition 2.6 Let $G = (V, E)$ be a directed graph. A *cycle* is a sequence $e_1 \cdots e_n$ of edges $e_i \in E$, $i = 1, \ldots, n$, such that

$$s(e_1) = t(e_n) \quad \text{and } t(e_i) = s(e_{i+1}) \text{ for } i = 1, \ldots, n-1 \ .$$

Definition 2.7 Let $G = (V, E)$ be a directed graph. The *period* $p(v)$ of a vertex $v \in V$ is the greatest common divisor of all cycles that start in v. If there exist no cycles, then we set $p(v) = \infty$.

The vertices of an irreducible graph have the same period. For that reason an irreducible graph with a vertex $v \in V$ whose period is $p(v) = 1$ is called *aperiodic*.

Lemma 2.8 *Let G be a road-coloured graph with a synchronising word. Then G is irreducible and aperiodic.*

Proof. For a proof we refer to [Lan03]. □

Let A, C be finite sets, ν a strictly positive probability measure on C and $\gamma\colon A \times C \to A$ a mapping. We construct a stochastic graph G_γ with vertices A; the set of edges is given by $E := \{\, (a,c) \in A \times C \colon a \in A, c \in C \,\}$. The mapping $\zeta\colon E \to C,\ (a,c) \mapsto c$ provides a natural colouring and $\omega\colon E \to [0,1],\ e \to \nu(\zeta(e))$ weights, such that (G_γ, ω) is a weighted graph with a road-colouring $\zeta\colon E \to C$.

Proposition 2.9 *Let $\gamma\colon (A \times C, \mu \otimes \nu) \to (A, \mu)$ be measure preserving, i.e. $\mu \otimes \nu(\gamma^{-1}(U)) = \nu(U)$ for all $U \subseteq C$. If G_γ has a synchronising word, then G_γ is irreducible.*

Proof. A proof may be found in [Lan03]. □

Theorem 2.10 *Let $J\colon (L^\infty(A,\mu), \varphi_\mu) \to (L^\infty(A,\mu) \otimes L^\infty(C,\nu), \varphi_\mu \otimes \varphi_\nu)$ be a random variable associated to the measure preserving transformation $\gamma\colon A \times C \to A$. The following assertions are equivalent:*

a) The graph G_γ possesses a synchronising word.

b) The random variable J is asymptotically complete.

Proof. For a proof we refer to [GKL06]. □

Regularity of a transition operator $T = P_\psi J$ is not sufficient to guarantee asymptotic completeness of the random variable J. In the light of theorem 2.10 we can give an easy counterexample.

Example 2.11 We put $A = \{1,2,3\}$ with measure $\mu = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $C = \{0,1\}$ and measure $\nu = (\frac{1}{2}, \frac{1}{2})$. Let the transition operator $T\colon A \to A$ be given by $T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$. In Figure 2.1 we present two graphs that both reproduce the transition matrix, but differ by their colouring. The first graph possesses a synchronising word, e.g. 00, while the second does not.

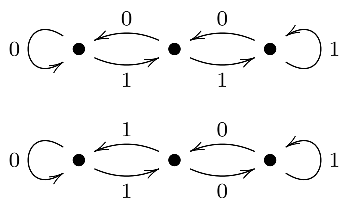


Figure 2.1: Roadcolouring with and without a synchronising word

Consequently, the random variable

$$\begin{aligned} J_1\colon (L^\infty(A,\mu),\varphi_\mu) \to &(L^\infty(A,\mu),\varphi_\mu)\otimes(L^\infty(C,\nu),\varphi_\nu),\\ &a\mapsto \begin{pmatrix} 1&0&0\\ 1&0&0\\ 0&1&0 \end{pmatrix}\cdot\begin{pmatrix} a_1\\ a_2\\ a_3 \end{pmatrix}\otimes\begin{pmatrix} 1\\ 0 \end{pmatrix}+\begin{pmatrix} 0&1&0\\ 0&0&1\\ 0&0&1 \end{pmatrix}\cdot\begin{pmatrix} a_1\\ a_2\\ a_3 \end{pmatrix}\otimes\begin{pmatrix} 0\\ 1 \end{pmatrix} \end{aligned}$$

is asymptotically complete, while the random variable

$$\begin{aligned} J_2\colon (L^\infty(A,\mu),\varphi_\mu) \to &(L^\infty(A,\mu),\varphi_\mu)\otimes(L^\infty(C,\nu),\varphi_\nu),\\ &a\mapsto \begin{pmatrix} 1&0&0\\ 0&0&1\\ 0&1&0 \end{pmatrix}\cdot\begin{pmatrix} a_1\\ a_2\\ a_3 \end{pmatrix}\otimes\begin{pmatrix} 1\\ 0 \end{pmatrix}+\begin{pmatrix} 0&1&0\\ 1&0&0\\ 0&0&1 \end{pmatrix}\cdot\begin{pmatrix} a_1\\ a_2\\ a_3 \end{pmatrix}\otimes\begin{pmatrix} 0\\ 1 \end{pmatrix} \end{aligned}$$

is not.

However, if T is an aperiodic and irreducible transition matrix and G_T a graph which represents T and can be road-coloured, then there exists a colouring with a synchronising word. This is a direct consequence of the road-colouring problem, which Thratman [Tra09] has proven is solvable. We state his result only in an abbreviated form, that serves all our needs.

Theorem 2.12 *Let G be an aperiodic and irreducible graph with constant outdegree. Then G possesses a road-colouring with a synchronising word.*

2.2 Classification of Transition matrices and Perron-Frobenius theorem

A $n \times n$-matrix $T \in M_n(\mathbb{R})$ is called stochastic if $T_{i,j} \in [0,1]$ for all $i, j = 1, \ldots, n$ and $\sum_{j=1}^{n} T_{i,j} = 1$. That means the rows of a stochastic matrix are probability measures on $\{1, \ldots, n\}$. With regard to their role for stochastic processes we refer to stochastic matrices as transition matrices, too. If additionally the columns are probability measures on $\{1, \ldots, n\}$ as well, the matrix T is called doubly stochastic. The set of stochastic as well as the set doubly stochastic matrices is a convex set. The extremal points of the former are those stochastic matrices that have only entries in $\{0, 1\}$, those of the latter are the permutation matrices. For details we refer to [AU82]. Let $X_t\colon (\Omega, \Sigma, \mu) \to \Gamma$, $t \in \mathbb{N}_0$, be a family of random variables with values in a finite set $\Gamma := \{1, \ldots, n\}$. By $\Sigma_{[0,t]}$ we denote the σ-algebra generated by the random variables $X_0, \ldots, X_t$.

Definition 2.13 We call $(X_n)_{n \in \mathbb{N}_0}$ a *Markov chain* with initial distribution $\lambda \in M(\Gamma)$ and transition matrix T if

i) $\mu(\{\omega \in \Omega\colon X_0(\omega) = i\}) = \lambda_i$,

ii) $\mu(X_{n+1} = i_{n+1} | X_n = i_n, \ldots, X_0 = i_0) = T_{i_n, i_{n+1}}$.

2.2.1 Irreducible and Aperiodic Markov Chains

Often the state space Γ of a Markov chain can be decomposed into disjoint subsets $\Gamma_k,\ k = 0, \ldots, d-1$, of positive measure that are invariant under T. Intuitively, the existence of a partition of invariant subsets characterises a dynamic that does not mingle the whole state space Γ properly. If the state space is indecomposable, the Markov chain is called irreducible.

Definition 2.14 A Markov chain is called *irreducible* if it is conform with one of the following equivalent conditions:

a) For every pair $i, j \in \Gamma$ there exist natural numbers $k_1, k_2 \in \mathbb{N}$, such that

$$\mu(\{\, \omega \in \Omega \colon X_{k_1}(\omega) = i,\ X_{k_2}(\omega) = j \,\}) > 0 \ .$$

b) All subsets $A \subset \Gamma$, such that $T^{-1}(A) := \{\, b \in \Gamma \colon \exists a \in A \colon T_{a,b} > 0 \,\} \subseteq A$, have measure $\mu(A) \in \{\, 0, 1 \,\}$.

Still, even if the Markov chain is irreducible, there might exist a partition $\Gamma_k\,,\ k = 0, \dots, d-1$, such that

$$T(\Gamma_k) := \{\, b \in \Gamma \colon \exists a \in \Gamma_k \colon T_{a,b} > 0 \,\} = \Gamma_{k-1},\ k = 0, \dots, d-1,$$

where the equality has to be understood modulo d. As an example consider any cyclic shift on a finite state space Γ.

Definition 2.15 An irreducible Markov chain is *periodic* with *period* d if there exists a partition $\{\, \Gamma_i \subset \Gamma \colon i = 0, \dots, d-1 \,\}$ such that

$$T(\Gamma_k) =_d \Gamma_{k-1},\ k = 1, \dots, d-1 \ .$$

In the standard literature on Markov processes this definition is slightly uncommon. There, the greatest common divisor of the length of all trajectories returning to a state $i \in \Gamma$ is assigned to i as period $p(i)$. By the independence of the choice of the state, the period characterises the Markov chain. Although both definitions are equivalent, we prefer the former being suited better for generalisation to non-commutative Markov processes. Note that if there is given an irreducible Markov process on Γ with transition matrix T which is d-periodic in the later sense, then for a fixed element $j \in \Gamma$

$$\Gamma_k := \left\{\, i \in \Gamma \colon \exists l \geq 1,\ T_{i,j}^{ld+k} > 0 \,\right\},\ k = 0, \dots, d-1,$$

defines a partition that satisfies the cyclic condition above.

Definition 2.16 An irreducible Markov chain is called *aperiodic* if $d = 1$ in the above definition.

2.2.2 Perron-Frobenius Theorem

Essentially, the Perron-Frobenius Theorem provides information about spectral properties of positive matrices or positive operators. In particular, as stochastic matrices are positive the theorem applies to them. Moreover the spectral radius of a stochastic matrix T is equal to 1. Consequently, spectral properties reveal the asymptotic behaviour of the powers of T, as the limit vanishes on eigenspaces E_λ with eigenvalue $|\lambda| < 1$. Therefore, the Perron-Frobenius Theorem has become a central part of the theory of Markov chains. For more details and proofs the book of E. Seneta [Sen81] is referenced.

Theorem 2.17 *Let $(X_n)_{n\in\mathbb{N}}$ be an irreducible Markov chain with transition matrix T and period d.*

i) The spectrum $\sigma(T)$ of T restricted to $\sigma(T) \cap \mathbb{T}$ is a finite subgroup of the torus group $\mathbb{T}$ with cardinality d.

ii) The right and left eigenspaces $E_s,\ s \in \sigma(T) \cap \mathbb{T}$ are one-dimensional.

iii) There exists a unique invariant positive probability measure $\pi_\infty \in \mathcal{M}(\Gamma)$.

Theorem 2.18 *Let $(X_n)_{n\in\mathbb{N}_0}$ be an irreducible Markov chain with transition matrix T. Then the following conditions are equivalent:*

a) T is aperiodic.

b) $\sigma(T) \cap \mathbb{T} = \{1\}$

In that case $\lim_{k\to\infty} T^k = P_{\mu_\infty}$, where $P_{\mu_\infty} = \begin{pmatrix} \pi_1 & \cdots & \pi_n \\ \vdots & & \vdots \\ \pi_1 & \cdots & \pi_n \end{pmatrix}$ with the invariant measure $\pi_\infty = (\pi_1, \ldots, \pi_n)$.

2.3 Classification of Transition Operators and Perron-Frobenius Theorem

2.3.1 Non-commutative Irreducible and Aperiodic Operators

A notion of irreducibility of a non-commutative Markov semigroup was introduced by F. Fagnola and R. Rebolledo in [FR02]. It very much translates definition 2.14.b) into the operator algebraic frame. Subsets of the classical state space are replaced by the corresponding subalgebra or the corresponding projections. The latter is suited for von Neumann algebras with their rich structure of projections.

Definition 2.19 Let $\mathcal{A}$ be a C^*-algebra and $T\colon \mathcal{A} \to \mathcal{A}$ a unital, completely positive operator. We call the operator T *irreducible* if:

a) There exists no non-trivial hereditary subalgebra $\mathcal{A}_0 \subsetneq \mathcal{A}$, such that $T(\mathcal{A}_0) \subseteq \mathcal{A}_0$.

If $\mathcal{A}$ is a W^*-algebra, a normal, unital and completely positive operator $T\colon \mathcal{A} \to \mathcal{A}$ is *irreducible* if T conforms one of the following equivalent conditions: The only projections $p \in \mathcal{A}$ that satisfy b), c) or d) are $p = 0$ and $p = \mathbb{1}$.

b) For all states $\varphi \in \mathcal{A}_*$ with $\operatorname{supp} \varphi \leq p$ follows $\operatorname{supp}(\varphi \circ T^n) \leq p$ for all $n \in \mathbb{N}_0$.

c) $T^n(p) \geq p$ for all $n \in \mathbb{N}_0$.

d) $T(p) \geq p$.

If $\mathcal{A}$ is a finite algebra, conditions a), b), c) and d) are equivalent.

Just like the concept of irreducibility one can generalise the concept of periodicity for operators on (finite-dimensional) von Neumann algebras [FB09]. In the usual algebraic approach the state space Γ is replaced by the corresponding W^*-algebra $\ell^\infty(\Gamma)$ and the partition $\Gamma_0, \ldots, \Gamma_{d-1}$ by the subalgebras $\ell^\infty(\Gamma_0), \ldots, \ell^\infty(\Gamma_{d-1})$.

Definition 2.20 Let $\mathcal{A}$ be a C^*-algebra and $T\colon \mathcal{A} \to \mathcal{A}$ a unital, completely positive and irreducible operator. It is *periodic* with *period* d if there exists a family of hereditary subalgebras $\mathcal{A}_0, \ldots, \mathcal{A}_{d-1}$ of $\mathcal{A}$, such that

$$T(\mathcal{A}_k) \subseteq_d A_{k-1} \ , \ k = 0, \ldots, d-1 \ .$$

In particular, if $\mathcal{A}$ is a von Neumann algebra, every hereditary von Neumann subalgebra can be written as $p\mathcal{A}p$ for a projection $p \in \mathcal{A}$. The operator T has period d if there exists a family of projections $\{ p_0, \ldots, p_{d-1} \} \subset \mathcal{A}$ with $\sum_{i=0}^{d-1} p_i = \mathbb{1}$, such that

$$T(p_k) =_d p_{k-1} \ , \ k = 0, \ldots, d-1 \ .$$

2.3.2 Irreducible Operators

In the following we will summarise a few properties of irreducible operators on von Neumann algebras. For this section we agree $\mathcal{A}$ to be a von Neumann algebra.

Lemma 2.21 *Every invariant state $\varphi \in \mathcal{S}(\mathcal{A})$ of an irreducible operator $T\colon \mathcal{A} \to \mathcal{A}$ is faithful.*

Proof. Let $\varphi \in \mathcal{S}(\mathcal{A})$ be an invariant state of the operator T. Then

$$\varphi \circ T^n(\mathbb{1} - \operatorname{supp} \varphi) = \varphi(\mathbb{1} - \operatorname{supp} \varphi) = 0 \ .$$

It follows that

$$T^n(\operatorname{supp} \varphi) \geq \operatorname{supp} \varphi \ .$$

Since T is irreducible, $\operatorname{supp} \varphi = \mathbb{1}$. □

Lemma 2.22 *Let $T\colon \mathcal{A} \to \mathcal{A}$ be a unital, completely positive operator with a faithful invariant state $\varphi \in \mathcal{S}$. The operator T is irreducible if and only if its fixed space $\mathcal{F}$ is $\mathcal{F} = \mathbb{C}\mathbb{1}$.*

Proof. Let the operator T be irreducible. Assume that $\mathcal{F} \supsetneq \mathbb{C}\mathbb{1}$. As $\mathcal{F}$ is a von Neumann algebra ([KN79], [Fri78]), there exists a projection $0 \neq p \in \mathcal{F}$. Of course,

$T(p) = p$. But this is a contradiction.
Now assume that T has a one-dimensional fixed space, but is not irreducible. Then there exists a projection $0 < p < 1$, such that $T(p) \le p$. Since in a von Neumann algebra monotone sequences converge pointwisely with respect to the strong operator topology, there exists a limit p_∞ of the sequence $T^n(p)$. Clearly, p_∞ is an element of the fixed space $\mathcal{F}$. The element p_∞ cannot be trivial, because otherwise

$$0 = \varphi(p_\infty) = \varphi(\lim_{n\to\infty} T^n(p)) = \varphi(p) \ne 0$$

would follow. This contradicts our assumption. □

Lemma 2.23 *Let $\mathcal{A}$ be $\mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$ and $T\colon \mathcal{A} \to \mathcal{A}$ a unital, completely positive operator with a one-dimensional fixed space $\mathcal{F}$ and a pure invariant state $\varphi \in \mathcal{S}(\mathcal{A})$. Then $\varphi \in \mathcal{S}(\mathcal{A})$ is an absorbing state, i.e.* $\lim_{n\to\infty} \theta \circ T(x) = \varphi(x)$ *for all $\theta \in \mathcal{S}(\mathcal{A})$ and $x \in \mathcal{A}$.*

Proof. Let $p \in \mathcal{A}$ be the support of φ. In the proof of Lemma 2.21 we have learnt that $T(p) \le p$, such that $(T^n(p))_{n\in\mathbb{N}_0}$ is a monotone and bounded sequence in $\mathcal{A}$. As such, it has a limit $p_\infty \in \mathcal{A}$ with respect to the strong operator topology, which is a fixed point. Since T has a one-dimensional fixed space

$$p_\infty = 0 \ .$$

Now consider $a \in \mathcal{A}$ decomposed into four terms

$$a = pap + p^\perp a p^\perp + pap^\perp + p^\perp a p \ .$$

The last three terms tend to 0 under $\theta \circ T^n$ for each $\theta \in \mathcal{S}(\mathcal{A})$. We exemplary show that for $pap^\perp$:

$$|\theta \circ T^n(pap^\perp)|^2 \le \theta \circ T^n(p)\theta \circ T^n(p^\perp a^* a p^\perp)$$

by the Cauchy-Schwartz inequality for states. Note that the first term can be written as

$$pap = \varphi(a)p \ .$$

Therefore $\lim T^n(pap) = \varphi(a)\mathbb{1}$ with respect to the σ-weak topology. □

As we have seen in Section 2.2.2, the analysis of the asymptotic behaviour of an irreducible Markov chain is considerably easier when the chain is aperiodic. By a simple modification of the transition matrix, we obtain a new Markov chain that, given the initial chain has a stationary distribution, has the same, yet is aperiodic. If $(X_n)_{n\in\mathbb{N}_0}$ is an irreducible Markov chain with transition matrix T and stationary distribution π, the chain $(Y_n)_{n\in\mathbb{N}_0}$ determined by the transition matrix $\overline{T} = \frac{1}{2}T + \frac{1}{2}\mathbb{1}$ is aperiodic and irreducible.

Proposition 2.24 *Let $\mathcal{A}$ be a finite algebra and $T\colon \mathcal{A} \to \mathcal{A}$ a completely positive, irreducible operator. The operator $\overline{T}\colon \mathcal{A} \to \mathcal{A}$ defined by $\overline{T}(x) := (1-\lambda)T(x) + \lambda x$ with $\lambda \in (0,1)$ is aperiodic and irreducible. If the operator T possesses an invariant state $\varphi \in \mathcal{S}(\mathcal{A})$, then it is invariant under $\overline{T}$ as well.*

Proof. First we show that $\overline{T}$ is irreducible. Assume the contrary is true. Then there exists a projection $p \in \mathcal{A}$, such that $\overline{T}(p) = (1-\lambda)T(p) + \lambda p \leq p$ with $p \notin \{\,0, \mathbb{1}\,\}$. From the inequality follows that $T(p) \leq p$, which is contrary to the irreducibility of the operator T.
Suppose $\overline{T}$ had period d, there would exist a family $\{p_i\}_{i\in I}$ of mutually orthogonal projections $p_i \in \mathcal{A}$, $i \in I$, with $\sum_{i\in I} p_i = \mathbb{1}$, such that

$$\overline{T}(p_i) = (1-\lambda)T(p_i) + \lambda p_i =_d p_{i+1}\ , \quad \text{for } i \in I\ .$$

As p_i and p_{i+1} are orthogonal this cannot happen. □

By $CP_{ir}(\mathcal{A},\mathcal{A})$ and $CP_{ap}(\mathcal{A},\mathcal{A})$ we denote the set of irreducible and the set of aperiodic, irreducible operators on $\mathcal{A}$; of course, $CP_{ap}(\mathcal{A},\mathcal{A}) \subset CP_{ir}(\mathcal{A},\mathcal{A})$.

Corollary 2.25 *Let $\mathcal{A}$ be a finite dimensional algebra. Then $CP_{ap}(\mathcal{A},\mathcal{A})$ is norm dense in $CP_{ir}(\mathcal{A},\mathcal{A})$.*

Example 2.26 Consider the unital, irreducible and completely positive operator

$$T\colon M_2 \to M_2,\ T(x) = \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right) x \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right) + \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right) x \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)\ .$$

The trace state φ with density matrix $\rho = \frac{1}{2}\mathbb{1}$ is invariant under T. The operator T has period $d = 2$, as the two projections $p_1 = \left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right)$ and $p_2 = \left(\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right)$ satisfy the cyclic condition:

$$T(p_1) = p_2 \quad \text{and} \quad T(p_2) = p_1 \ .$$

In place of T we consider the operator $S\colon M_2 \to M_2,\ S(x) = \frac{1}{2}T(x) + \frac{1}{2}x$. The operator S is aperiodic and irreducible.

Ergodic Properties

At this point we will return to section 1.2.14 shortly. A question that almost suggests itself to ask, is about the relation between asymptotic completeness of the random variable J we have introduced in section 1.2.14 and ergodic properties of the transition operator T_ψ. As to the most fundamental ergodic property the following is true.

Definition 2.27 An injective *-homomorphism $J\colon \mathcal{A} \to \mathcal{A} \otimes \mathcal{C}$ is called *irreducible* if all projections $p \in \mathcal{A}$ with $J(p) \leq p \otimes \mathbb{1}$ satisfy $p \in \{ 0, \mathbb{1} \}$.

For proofs of the following three propositions we refer the reader to [GKL06].

Proposition 2.28 *Let ψ be a faithful state on $\mathcal{C}$. Then the random variable J is irreducible if and only if T_ψ is irreducible.*

Proposition 2.29 *Let ψ be a faithful state on $\mathcal{C}$ and $J\colon \mathcal{A} \to \mathcal{A} \otimes \mathcal{C}$ an irreducible injective *-homomorphism. Then T_ψ is irreducible and possesses a unique invariant state $\varphi \in \mathcal{S}(\mathcal{A})$, which is faithful.*

Proposition 2.30 *If $J\colon (\mathcal{A}, \varphi) \to (\mathcal{A} \otimes \mathcal{C}, \varphi \otimes \psi)$ is an asymptotically complete random variable, then φ is absorbing, i.e.*

$$\lim_{n\to\infty} T_\psi^n(a) = \varphi(a)\mathbb{1} \ .$$

However, if J is irreducible and T_ψ possesses an absorbing state $\varphi \in \mathcal{S}(\mathcal{A})$, J does not necessarily have to be asymptotically complete. Example 2.11 provides a classical counterexample.

Proposition 2.31 *Let A and C be finite sets and $(X_n)_{n\in\mathbb{N}}$ an aperiodic, irreducible Markov chain on A with transition matrix T and invariant measure μ. Let further ν be the uniform distribution $\nu(c) = \frac{1}{|C|}$ on C. Then the following assertions are equivalent.*

a) There exists a random variable

$$J_0 \colon (L^\infty(A,\mu), \varphi_\mu) \to (L^\infty(A,\mu) \otimes L^\infty(C,\nu), \varphi_\mu \otimes \psi_\nu)$$

with $P_{\psi_\nu} J_0 = T$.

b) There exists an asymptotically complete random variable

$$J \colon (L^\infty(A,\mu), \varphi_\mu) \to (L^\infty(A,\mu) \otimes L^\infty(C,\nu), \varphi_\mu \otimes \psi_\nu)$$

with $P_\psi J = T$.

Proof. Let $\gamma_0 \colon A \times C \to A$ be the surjective measure preserving transformation corresponding to J_0. The graph G_{γ_0} generated by γ_0 is aperiodic and irreducible. By Theorem 2.12 we may modify the colouring γ_0, such that we obtain a new colouring $\gamma \colon A \times C \to A$ that has a synchronising word. The associated random variable $J \colon (L^\infty(A,\mu), \varphi_\mu) \to (L^\infty(A,\mu) \otimes L^\infty(C,\nu), \varphi_\mu \otimes \psi_\nu)$ is asymptotically complete. Note that in general recolouring a stochastic graph does not yield a stochastic graph. However, as ν is the uniform distribution on C, the graph G_γ is stochastic again. □

2.3.3 Non-commutative Perron-Frobenius Theorem

Naturally assertions valid for matrices pass through generalisation to further structures. Indeed, the first generalisation to positive operators on ordered linear spaces appeared in 1950 (e.g. [KR50]). In the 1980's U. Groh [Gro81] as well as D. Evans and R. Høegh-Krohn [EHK78] stated and proved corresponding theorems for positive operators on C^*-algebras. Introducing the above notion of periodicity

F.Fagnola and R.P. Bidot [FB09] published a different proof of the same results for finite dimensional C^*-algebras more in line with the classical approach.

Theorem 2.32 *Let $\mathcal{A}$ be a finite dimensional algebra and $T\colon \mathcal{A} \to \mathcal{A}$ a unital, irreducible and completely positive operator with period d.*

i) The spectrum $\sigma(T)$ of T restricted to $\sigma(T) \cap \mathbb{T}$ is a finite subgroup of the torus group $\mathbb{T}$ with cardinality d.

ii) The eigenspaces E_s, $s \in \sigma(T) \cap \mathbb{T}$, are one-dimensional.

iii) There exists a unique invariant state $\varphi_\infty \in \mathcal{S}(\mathcal{A})$.

Theorem 2.33 *Let $\mathcal{A}$ be a finite dimensional algebra and $T\colon \mathcal{A} \to \mathcal{A}$ a unital, irreducible and completely positive operator. Then the following conditions are equivalent:*

a) T is aperiodic.

b) $\sigma(T) \cap \mathbb{T} = \{ 1 \}$.

In that case $\lim_k T^k = E_{\varphi_\infty}$, where $E_{\varphi_\infty}(x) := \varphi_\infty(x)\mathbb{1}$ with the invariant state $\varphi_\infty \in \mathcal{S}(\mathcal{A})$.

CHAPTER 3

COUPLING FROM THE PAST

Let A be a finite state space and $(X_n)_{n \in \mathbb{N}_0}$ an aperiodic and irreducible Markov chain with unknown stationary distribution π. We would like to produce random samples distributed according to π. One way of attacking this problem is by means of the Perron-Frobenius Theorem (compare Section 2.2.2): It states for an aperiodic and irreducible Markov chain $(X_n)_{n\in\mathbb{N}_0}$ with transition matrix T

$$\|\lambda \circ T^n - \pi\|_{var} \to 0 \quad \text{for } n \to \infty \ ,$$

for every initial distribution λ, where $\| \cdot \|_{var}$ is the variation distance. Even more, the rate of convergence is exponential, $\|\lambda T^n - \pi\|_{var} = o(n^{-\alpha})$ for some $\alpha > 0$. We are able to produce samples with distribution arbitrarily close to π (in variation distance). Unfortunately, this is not sufficient. We would like the samples to be exactly distributed according to π, and we do not want to wait infinitely long almost surely.

By a both astonishingly simple and effective modification James G. Propp and David B. Wilson [PW96] created an algorithm that is able to produce exact random samples in finite time almost surely. Being also based on coupling methods and hinting at the modification, the algorithm became known as ‘Coupling from the

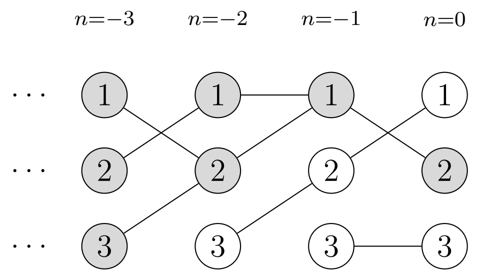


Figure 3.1: Coalescing trajectories

Past'. We will use the abbreviation CFTP hence. Next we illustrate their algorithm by an example. A further analysis will follow later on.

Example 3.1 Consider the transition matrix

$$T = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

on the state space $A := \{1, 2, 3\}$. Together with some initial distribution $\lambda \in \mathcal{M}(A)$, the probability measures on A, this provides us with a Markov chain. Let π with $\pi \circ T = \pi$ be the stationary distribution. To sample exactly from this chain, what we have to do according to CFTP is:

Step 1: Start in the presence at time $t = 0$ and step back once into the past. To each $a \in A$ find a successor in accordance with the transition matrix T. That means for state 1 choose a successor from $\{1, 2\}$ with probabilities $\mathbb{P}(1 \to 1) = T_{1,1} = \frac{1}{3}$ and $\mathbb{P}(1 \to 2) = T_{1,2} = \frac{2}{3}$, for state 2 a successor from $\{1, 3\}$ with probabilities $\mathbb{P}(2 \to 1) = T_{2,1} = \frac{1}{3}$ and $\mathbb{P}(2 \to 3) = T_{2,3} = \frac{2}{3}$ and for state 3 a successor from $\{2, 3\}$ with probabilities $\mathbb{P}(3 \to 2) = T_{3,2} = \frac{1}{3}$ and $\mathbb{P}(3 \to 3) = T_{3,3} = \frac{2}{3}$. In Figure 3.1 we have picked the transitions $1 \to 2$, $2 \to 1$ and $3 \to 3$.

Step 2: Then step back again and find successors to each $a \in A$.

Step 3: Repeat this until you obtain three trajectories, exactly one starting from each $a \in A$ at the same time $-n$ in the past, such that all of them coalesce in one state $a_0 \in A$ in the presence. The coalescence shown in Figure 3.1 happened after three steps. We have highlighted the states along the coalescing trajectories by grey colouring.

Step 4: Put out a_0. Accept a_0 as random sample. According to Figure 3.1 our random sample would be the state 2.

J. G. Propp and D. B. Wilson have shown that for every aperiodic and irreducible Markov chain this procedure will produce an output in finite time almost surely and that the output will be distributed as π. As in the above example the chain was chosen aperiodic and irreducible, our output would have been an exact random sample.
A justified objection often raised at this point concerns the necessity of running CFTP from the past. Would it not be the same to run CFTP into the future and let the trajectories coalesce at time $n > 0$? The answer is no, as the following counterexample shows.

Example 3.2 Again we choose the state space $A := \{1, 2, 3\}$ and consider the transition matrix

$$T = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

with invariant measure $\pi = (\frac{1}{10}, \frac{3}{10}, \frac{6}{10})$. Note that there is only one possible transition to state 1; that from 2 to 1 with probability $T_{21} = \frac{1}{3}$. Therefore, every trajectory which passes through 1 at time $t = n$ had to be in state 2 at $t = n-1$. Consequently, this kind of forward algorithm will never put out state 1, but always will already have stopped at state 2. Thus, the output will be distributed according to $(0, \nu_2, \nu_3)$, $0 \leq \nu_2, \nu_3 \leq 1$, $\sum_{i=1}^{3} \nu_i = 1$. Although $\pi = (\frac{1}{10}, \frac{3}{10}, \frac{6}{10})$.

3.1 Composition into the Past

In this section we set up a notation to describe CFTP formally and present a first proof. As we will see randomly choosing a successor to each $a \in A$ may as well be

regarded as randomly choosing a function onto A - only our viewpoint has changed a little, from state space to function space. On that space we will define a process associated to CFTP.
Let A be a finite state space and $(X_n)_{n\in\mathbb{N}_0}$ a Markov chain with transition matrix T. We will denote by F_A the space of all functions from A to A, $F_A := \{f : A \to A\}$. Note that $(F_A, \circ)$ with $\circ : F_A \times F_A \to F_A,\ (g, f) \mapsto f \circ g$ is a semigroup. Since A is finite, so is F_A; the choice of a σ-algebra Σ for F_A is therefore the power set $\mathcal{P}(F_A)$. We will omit the specification of the σ-algebra of F_A in the following. On Σ we define a measure ν by setting

$$\nu(f) = \prod_{i\in A} T_{i,f(i)},\ \ f \in\ F_A\,,$$

on a singleton $\{\,f\,\} \in \Sigma$ and extend it to the entire σ-algebra. Hereby, we give every mapping $f \in F_B$ a measure that corresponds to the product of the weights $T_{i,f(i)}$ of the transitions $i \mapsto f(i)$ it is formed of. This construction plays a key role in the following section. As we will see later on, it is universal for every (aperiodic and irreducible) Markov chain.

Proposition 3.3 ν is a probability measure on Σ.

Proof. Only the assertion $\nu(F_A) = 1$ needs some additional examination. For this proof we extend the notation that we have introduced for the space of functions from A to A. If B is another finite set, we denote the set $\{\,f\colon B \to A\,\}$ of all functions from B to A by F_B. We identify the space of functions F_A with the set $F_{\{\,i_0\,\}} \times F_{A\setminus\{\,i_0\,\}}$ by the following bijection

$$F_A \ni f \mapsto (f|_{\{\,i_o\,\}}, f|_{A\setminus\{\,i_0\,\}}) \in F_{\{\,i_0\,\}} \times F_{A\setminus\{\,i_0\,\}}\ .$$

By definition of the measure ν we obtain

$$\nu(F_A) = \sum_{f\in F_A} \nu(\{\,f\,\}) = \sum_{f\in F_A} \prod_{i\in A} T_{i,f(i)} = \sum_{f\in F_A} T_{i_0,f(i_0)} \prod_{i\in A\setminus\{i_0\}} T_{i,f(i)}\ .$$

Using the above identification $F_A \simeq F_{\{\,i_0\,\}} \times F_{A\setminus\{\,i_0\,\}}$ we split the sum into a double sum:

$$\begin{aligned}\sum_{f\in F_A} T_{i_0,f(i_0)} \prod_{i\in A\setminus\{i_0\}} T_{i,f(i)} &= \sum_{g\in F_{\{i_0\}}} \sum_{f\in F_{A\setminus\{i_0\}}} T_{i_0,g(i_0)} \prod_{i\in A\setminus\{i_0\}} T_{i,f(i)} \\ &= \left(\sum_{g\in F_{\{i_0\}}} T_{i_0,g(i_0)}\right)\left(\sum_{f\in F_{A\setminus\{i_0\}}} \prod_{i\in A\setminus\{i_0\}} T_{i,f(i)}\right) \ .\end{aligned}$$

Since T is a stochastic matrix, we have $\sum_{g\in F_{\{i_0\}}} T_{i_0,g(i_0)} = 1$, such that

$$\nu(F_A) = \sum_{f\in F_A\setminus\{i_o\}} \prod_{i\in A\setminus\{\, i_0\,\}} T_{i,f(i)} \ .$$

Inductively splitting the sum and factoring out ones yields

$$\nu(F_A) = 1 \ . \qquad \square$$

Now we are provided with a probability space (F_A, Σ, ν). We set $\Omega := \{\, F_A \,\}^{\mathbb{Z}}$, $\hat{\nu} := \otimes\nu$, the product measure of ν, and consider the product space $(\Omega, \otimes\Sigma, \hat{\nu})$ with the product σ-algebra $\otimes\Sigma$.

Definition 3.4 The *backward composition process* $(\hat{B}_n)_{n\in\mathbb{N}}$ is given by

$$\hat{B}_n : \Omega \to F_A, \quad \omega = (\cdots, f_{-1}, f_0, f_1, \cdots) \mapsto f_{-1} \circ \cdots \circ f_{-n}$$

and dually the *forward composition process* $(\hat{F}_n)_{n\in\mathbb{N}}$, by

$$\hat{F}_n : \Omega \to F_A, \quad \omega = (\cdots, f_{-1}, f_0, f_1, \cdots) \mapsto f_{n-1} \circ \cdots \circ f_0.$$

By restriction to the finite product spaces $\prod_{i=-n}^{-1} F_A$ and $\prod_{i=0}^{n-1} F_A$ the random variables $\hat{B}_n$ and $\hat{F}_n$ can be regarded as a mapping from the finite product space $\{\, F_A \,\}^n$ to F_A. Whenever we refer to the finite composition, we will denote this by B and F.

If the meaning of the backward composition process is not clear yet, it will become soon. At first, however, we will discover a few of its properties.

Lemma 3.5 *For fixed $\omega \in \Omega$ assume there exists a natural number $n \in \mathbb{N}$, such that $\hat{B}_n(\omega) \equiv constant$, i.e., all states are mapped to a single state. Then for all $f \in F_A$*

$$\hat{B}_n(\omega) \circ f = \hat{B}_n(\omega) \quad \text{and} \quad f \circ \hat{B}_n(\omega) \equiv constant \ .$$

The analogous assertion is true for the forward composition $\hat{F}_n$.

Proof. The proof is obvious. □

Proposition 3.6 *$(\hat{B}_n)_{n\in\mathbb{N}}$ and $(\hat{F}_n)_{n\in\mathbb{N}}$ are Markov processes.*

Proof. We have to show that for given functions $g_1, \ldots, g_n \in F_A$

$$\hat{\nu}(\hat{B}_n = g_n | \hat{B}_{n-1} = g_{n-1}, \ldots, \hat{B}_1 = g_1) = \hat{\nu}(\hat{B}_n = g_n | \hat{B}_{n-1} = g_{n-1}) \ .$$

Note that $\hat{B}_n(\omega) = \hat{B}_{n-1}(\omega) \circ f_{-n}$ for every $\omega = (\cdots, f_{-1}, f_0, f_1, \cdots) \in \Omega$. Explicitly calculating the conditional probabilities yields

$$\begin{aligned}
\hat{\nu}(\hat{B}_n =& g_n | \hat{B}_{n-1} = g_{n-1}, \ldots, \hat{B}_1 = g_1) = \frac{\hat{\nu}([\hat{B}_n = g_n] \cap \cdots \cap [\hat{B}_1 = g_1])}{\hat{\nu}([\hat{B}_{n-1} = g_{n-1}] \cap \cdots \cap [\hat{B}_1 = g_1])} \\
=& \frac{\hat{\nu}(\{\, \omega \in \Omega \colon \omega_{-n} = f,\ g_n = g_{n-1} \circ f \,\} \cap [\hat{B}_{n-1} = g_{n-1}] \cap \cdots \cap [\hat{B}_1 = g_1])}{\hat{\nu}([\hat{B}_{n-1} = g_{n-1}] \cap \cdots \cap [\hat{B}_1 = g_1])} \\
=& \frac{\hat{\nu}(\{\, \omega \in \Omega \colon \omega_{-n} = f,\ g_n = g_{n-1} \circ f \,\})\hat{\nu}([\hat{B}_{n-1} = g_{n-1}] \cap \cdots \cap [\hat{B}_1 = g_1])}{\hat{\nu}([\hat{B}_{n-1} = g_{n-1}] \cap \cdots \cap [\hat{B}_1 = g_1])} \\
=& \hat{\nu}(\{\, \omega \in \Omega \colon \omega_{-n} = f,\ g_n = g_{n-1} \circ f \,\}) \\
=& \frac{\hat{\nu}(\{\, \omega \in \Omega \colon \omega_{-n} = f,\ g_n = g_{n-1} \circ f \,\})\hat{\nu}([B_{n-1} = g_{n-1}])}{\hat{\nu}([B_{n-1} = g_{n-1}])} \\
=& \frac{\hat{\nu}(\{\, \omega \in \Omega \colon \omega_{-n} = f,\ g_n = g_{n-1} \circ f \,\} \cap [B_{n-1} = g_{n-1}])}{\hat{\nu}([B_{n-1} = g_{n-1}])} = \\
=& \hat{\nu}(\hat{B}_n = g_n | \hat{B}_{n-1} = g_{n-1}) \ .
\end{aligned}$$

□

By A_∞ we denote the set $A_\infty := \Big\{ \omega \in \Omega \colon \exists n \in \mathbb{N} \colon \hat{B}_n(\omega) \equiv constant \Big\}$.

Lemma 3.7 *If $(X_n)_{n\in\mathbb{N}}$ is an aperiodic and irreducible Markov chain, then*

$$\hat{\nu}\left(A_\infty\right) > 0 \ .$$

Proof. The main idea of the proof is to choose for each $a \in A$ trajectories of the Markov chain starting in a and meeting in a single state after finitely many time steps. We will construct a finite sequence of functions of F_A originating from these trajectories whose backward composition is a constant function.
As the Markov chain is aperiodic and irreducible there exists some $n \in \mathbb{N}$, such that the n-th power T^n of the transition matrix has strictly positive entries. Thus, if we choose a state $b \in A$, we find for each $a \in A$ trajectories of the Markov chain that start in a and meet in b at finite time n. Note that these trajectories are formed of transitions that have non zero probability.

Now, we choose a state $b \in A$, thought as a target state, and for each $a \in A$ one trajectory $(\theta_n^a)_{n\in\mathbb{N}}$ of the Markov chain $(X_n)_{n\in\mathbb{N}}$ with $\theta_0^a = a$, such that:

1. Trajectories that have met at a time will evolve identically afterwards, i.e., if $\theta_k^a = \theta_k^c$ for $a, c \in A$, $a \neq c$ and some $k \in \mathbb{Z}$, then $\theta_l^a = \theta_l^c$ for all $l \geq k$;
2. All trajectories meet in b at finite time $t = n$;

The first requirement is easily met, as we are free to set two trajectories $(\theta_n^a))_{n\in\mathbb{N}}$ and $(\theta_n^{a'}))_{n\in\mathbb{N}}$ with $a \neq a' \in A$ equal to the first, after they have met. The second demand can be satisfied according to our comment on powers T^n of the matrix T.
We are now using the trajectories to construct a sequence of functions. We define for $k = 1, \dots, n$

$$f_{-k}: \ A \to A, \ c \mapsto \begin{cases} \theta(a, k+1) & \text{if there exists a } a \in A, \text{ with } c = \theta(a,k)\,, \\ d & \text{else, for an arbitrary } d \text{ with } \ T_{c,d} > 0. \end{cases}$$

The measures $\nu(\{ f_{-k} \}$ of the functions f_{-k}, $k = 1, \dots, n$ as elements of the function space F_A are strictly positive by their definition

$$\nu(\{ f_k \}) = \prod_{a\in A} T_{a, f_k(a)} > 0 \ .$$

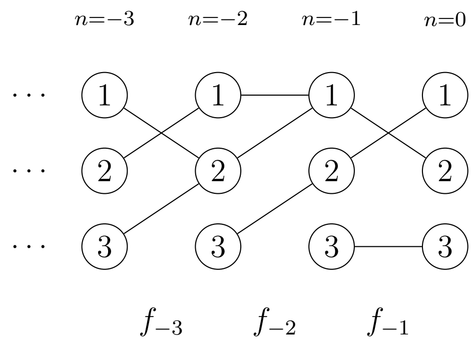


Figure 3.2: b-composition

Hence, the measure of the corresponding cylinder set

$$C := \{\, \omega \in \Omega \colon \omega_{-n} = f_{-n}, \dots, \omega_{-1} = f_{-1} \,\}$$

is strictly positive as well: $\hat{\nu}(C) = \prod_{k=1}^{n} \nu(\{\, f_{-k} \,\}) > 0$. Originating from the trajectories $\{\, (\theta_n^a)_{n\in\mathbb{N}} \colon a \in A \,\}$, the sequence $(f_{-n}, \cdots, f_{-1})$ has the property

$$B(f_{-n}, \cdots, f_{-1}) = f_{-1} \circ \cdots \circ f_{-n} = b \;.$$

□

As we will see even $\hat{\nu}(A_\infty) = 1$ is true.

Now we return to the question how the process $(\hat{B}_n)_{n\in\mathbb{N}}$ is related to CFTP. Again, we refer to Example 3.1 to illustrate this connection. In the following we will use a notation for functions by arrays with square brackets, which is borrowed from the common notation of permutations. If $f \colon B \to B$ is a mapping on some finite set B with $|B| = n$, we write

$$f = \begin{bmatrix} b_1 & f(b_1) \\ \vdots & \vdots \\ b_n & f(b_n) \end{bmatrix} \;.$$

Example 3.8 In Figure 3.2 we recognise those three steps that have led to coalescence of the three trajectories before. We consider the transitions between the states at each of the time steps as functions on A. To indicate them we have put symbols f_{-1}, f_{-2} and f_{-3} under the transitions. In the above notation those functions are given by

$$f_{-3} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}, \; f_{-2} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 2 \end{bmatrix} \text{ and } f_{-1} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 3 \end{bmatrix} .$$

Calculating $\hat{B}_3(f_{-3}, f_{-2}, f_{-1})$ yields

$$\hat{B}_3(f_{-3}, f_{-2}, f_{-1}) = \begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 3 & 2 \end{bmatrix} .$$

In this new setting, the procedure, we have described algorithmically before, reduces to observing the backward composition process $(\hat{B}_n)_{n\in\mathbb{N}}$ until $\hat{B}_n$ returns a constant function for some $n \in \mathbb{N}$. In our example this was fulfilled for $n = 3$, and we received $\hat{B}_3 \equiv 2$.

Remark 3.9 We are dealing with a stopping time

$$\mathcal{T} : \Omega \to \mathbb{N}_0, \quad \omega \mapsto \min\left\{ n \in \mathbb{N}_0 : \hat{B}_n(\omega) \equiv \textit{constant} \right\} .$$

So on Ω the algorithm turns out to be nothing else than the backward composition process stopped at the stopping time $\mathcal{T}$.

We conclude this section with a proof of Propp and Wilson's Theorem, reformulated in this notation. From Lemma 3.7 we know $\hat{\nu}(A_\infty) > 0$. Therefore, the limit $\lim_{n\to\infty} \hat{B}_n(\omega)$ exists for all $\omega \in A_\infty$ without further reference to any topology. We introduce the symbol $\hat{B}_\infty(\omega)$ for the limit function on A_∞.

Theorem 3.10 *Let $(X_n)_{n\in\mathbb{N}_0}$ be an aperiodic and irreducible Markov chain with stationary distribution π on a finite state space A. Then $\hat{\nu}\left(A_\infty\right) = 1$ and $\hat{B}_\infty$ is a random variable $\hat{\nu}$-almost everywhere distributed according to π.*

Proof. Firstly, we are going to proof that A_∞ is invariant under shift-transformation and conclude by a 0-1 argument that $\hat{\nu}(A_\infty) = 1$. Secondly, we show that A_∞ can be divided into $|A|$ disjoint subsets A_j of measure $\hat{\nu}(A_j) = \pi_j$ with $A_j := \Big\{ \omega \in A_\infty : \hat{B}_\infty(\omega) \equiv j \Big\}$.
The shift-transformation σ is given by

$$\sigma : \Omega \to \Omega,\ (\sigma(\omega))_n = f_{n-1}.$$

Suppose $\omega \in A_\infty$, then there exists $n \in \mathbb{N}$, such that $\hat{B}_n(\omega) = f_0 \circ \cdots \circ f_{-n} \equiv$ *constant*. We obtain

$$\begin{aligned} \hat{B}_n(\sigma^{-1}(\omega)) &= f_0 \circ f_{-1} \circ \cdots \circ f_{-n+1} \quad \text{and} \\ \hat{B}_{n+1}(\sigma^{-1}(\omega)) &= f_0 \circ f_{-1} \circ \cdots \circ f_{-n} = f_0 \circ \hat{B}_n(\omega) \equiv constant\ . \end{aligned}$$

Therefore $\sigma^{-1}(\omega) \in A_\infty$, and A_∞ is invariant under σ.
As σ is ergodic either $\hat{\nu}(A_\infty) = 0$ or $\hat{\nu}(A_\infty) = 1$ [Pet97]. From Lemma 3.7 it follows that $\hat{\nu}(A_\infty) = 1$. This proves the first part of the theorem. We emphasise the crucial role of Lemma 3.7: We can only hope CFTP to terminate if we have guaranteed the existence of a subset of A_∞ with non zero measure.
In order to show $\hat{\nu}(A_j) = \pi_j$ we define sets

$$A_{ij}^{(n)} := \Big\{ \omega \in A_\infty : \hat{B}_n(\omega)(i) = j \Big\}.$$

In general, these sets have no particular order and are neither included in any A_j nor do they contain any. Note, however, that they have two useful properties:

$$\begin{aligned} &1.)\quad \forall\, \omega \in A_j,\, \exists\, n_0 \in \mathbb{N} : \ \forall i \in A\ \forall\, n \geq n_0\, \omega \in A_{ij}^{(n)}\ ; \\ &2.)\quad \forall\, \omega \in \mathcal{A}_j,\ k \neq j\ \exists\, n_0 : \ \forall i \in\ A\,\forall\, n \geq n_0\, \omega \notin\ A_{ik}^{(n)}\ . \end{aligned}$$

The intersections $S_{ij}^{(m)} := \bigcap_{n \geq m} A_{ij}^{(n)}$ are subsets of A_j for all $i \in A$:

$$S_{ij}^{(m)} := \bigcap_{n \geq m} A_{ij}^{(n)} = \Big\{ \omega \in A_\infty : \hat{B}_n(\omega)(i) = j\ \forall\ n \geq m \Big\} \subseteq A_j\ ,$$

It follows that

$$\bigcup_{m=0}^{\infty} S_{ij}^{(m)} = A_j.$$

As the sets $S_{ij}^{(m)}$ are monotonously increasing, taking the measure of both sides yields

$$\hat{\nu}(A_j) = \hat{\nu}\Big(\bigcup_{m=0}^{\infty} S_{ij}^{(m)}\Big) = \lim_{n\to\infty} \hat{\nu}(S_{ij}^{(n)}) \ .$$

The last term can be estimated by

$$\lim_{n\to\infty} \hat{\nu}(S_{ij}^{(n)}) \leq \lim_{n\to\infty} \hat{\nu}(A_{ij}^{(n)}) \ .$$

Since for an aperiodic and irreducible Markov chain, the n-step transition probabilities, i.e., $\hat{\nu}(A_{ij}^{(n)})$, converge to the stationary distribution by Theorem 2.18, we obtain $\hat{\nu}(A_j) \leq \pi_j$. From this and

$$\sum_{j\in A} \hat{\nu}(A_j) = \hat{\nu}(A_\infty) = 1 = \sum_{j\in A} \pi_j$$

we conclude that $\hat{\nu}(A_j) = \pi_j$. □

3.2 Semigroups of Functions

In this section we would like to show that the semigroup F_A used in the last proof is larger than actually necessary. Note that if A has cardinality $|A| = n$, $|F_A| = n^n$ already. By demanding a certain property, we will be able to choose a (smaller) sub-semigroup G of F_A which is sufficient for the implementation of CFTP. It is a reinterpretation of what is called an update function in the literature. How those two perspectives translate into each other will be discussed in the following.

Definition 3.11 Let $G \subseteq F_A$ be a sub-semigroup and μ a measure on G. We call (G, μ) a *global coupling* of the Markov chain $(X_n)_{n\in\mathbb{N}_0}$ if $\mu(\{f \in G : f(i) = j\}) = T_{i,j}$ for all $i, j \in A$.

Proposition 3.12 *Let $(X_n)_{n\in\mathbb{N}_0}$ be an aperiodic and irreducible Markov chain. Then (F_A,ν) is a global coupling.*

Proof. We define $F^{ij} := \{f \in F_A : f(i) = j\}$ for given $i,j \in A$. Recall the definition of the measure ν on F_A:

$$\nu(\{f\}) = \prod_{i\in A} T_{i,f(i)} \ .$$

This gives

$$\begin{aligned}\nu(F^{ij}) &= \sum_{f\in F^{ij}} \hat{\nu}(\{f\}) = \sum_{f\in F^{ij}} \prod_{k\in A} T_{k,f(k)} = T_{i,j} \sum_{f\in F^{ij}} \prod_{k\in A\setminus\{i\}} T_{k,f(k)} \\ &= T_{i,j} \sum_{f\in F_{A\setminus\{i\}}} \prod_{k\in A\setminus\{i\}} T_{k,f(k)} = T_{i,j} \cdot 1 = T_{i,j} \ .\end{aligned}$$

For the second last equality we have been splitting the sum into a double sum successively as in the proof of Lemma 3.5, and factored out ones. □

Definition 3.13 We call a global coupling (G,μ) *synchronising* if there exists a set of functions $\{ f_0,\dots,f_n \} \subseteq G$, such that $\mu(f_i) \neq 0$ for all $i = 0,\dots,n$ and $f_n \circ \cdots \circ f_0 \equiv$ *constant*.

Remark 3.14 G is only synchronising if and only if there exists an element $g \in G$ that is a constant function. We will call a constant function g synchronising as well if there is no chance for misinterpretation.

Reinterpreting Lemma 3.5 in algebraic terms yields:

Lemma 3.15 *Let $G \subseteq F_A$ be a semigroup and $G_{sync} \subset G$ the sub-semigroup of constant functions. Then G_{sync} is an ideal, and $f \circ g = f$ holds for all $f \in G_{sync}$ and $g \in G$.*

Proposition 3.16 *Let $(X_n)_{n\in\mathbb{N}_0}$ be an aperiodic and irreducible Markov chain. Then (F_A,ν) is synchronising.*

Proof. Recall Lemma 3.7. □

The idea of Definition 3.13 is based on the importance of Lemma 3.7 for proving Theorem 3.10. Without the synchronising property Theorem 3.10 would not be true.

Example 3.17 On $A := \{1, 2, 3\}$ we consider the Markov chain $(X_n)_{n\mathbb{N}_0}$ with transition matrix $T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

a) The semigroup $G_1 := \left\{ \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 2 \end{bmatrix} \right\}$ with $\mu(\begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 3 & 2 \end{bmatrix}) = 1$ is neither synchronising nor a global coupling of the Markov chain associated with T.

b) The semigroup G_2 generated by $f_0 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 3 \end{bmatrix}$ and $f_1 = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 2 \end{bmatrix}$ with $\mu(f_0) = \frac{1}{2}$, $\mu(f_1) = \frac{1}{2}$ is a global coupling of the Markov chain, but not synchronising. It is a global coupling, because an easy calculation shows $\mu(\{ f \in G_2 \colon f(i) = j \}) = T_{ij}$ for all $i, j \in A$, but not synchronising, as the generating functions are bijections.

c) The semigroup G_3 generated by $f_0 = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$ and $f_1 = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 3 \end{bmatrix}$ with $\mu(f_0) = \frac{1}{2}$, $\mu(f_1) = \frac{1}{2}$ is a synchronising global coupling of the Markov chain. As we will return to this semigroup in Example 3.21 and will state it explicitly then, we postpone the verification.

It suggests itself to ask under which conditions a Markov chain $(X_n)_{n\in\mathbb{N}_0}$ has a synchronising global coupling. From Proposition 3.16 and 3.12 we find:

Corollary 3.18 *Let $(X_n)_{n\in\mathbb{N}_0}$ be an aperiodic and irreducible Markov chain on a finite space A. Then (F_A, ν) is a synchronising global coupling for $(X_n)_{n\in\mathbb{N}_0}$.*

If $G \subseteq F_A$ is a semigroup with a measure μ, we consider the product space $(\Omega_G, \Sigma_G, \hat{\mu})$ with $\Omega_G := G^{\mathbb{Z}}$, the product σ-algebra $\Sigma_G := \otimes_{\mathbb{Z}}\Sigma$ and the product measure $\hat{\mu} := \otimes_{\mathbb{Z}}\mu$. Similar to the composition processes we constructed in the last section for the semigroup $G = F_A$, we define processes on Ω_G with values in the semigroup $G \subset F_A$.

Definition 3.19 The *backward composition process* $(\hat{B}_n)_{n\in\mathbb{N}}$ is given by

$$\hat{B}_n : \Omega_G \to G, \quad \omega = (\cdots, f_{-1}, f_0, f_1, \cdots) \mapsto f_0 \circ \cdots \circ f_{-n}$$

and dually the *forward composition process* $(\hat{F}_n)_{n\in\mathbb{N}}$, by

$$\hat{F}_n : \Omega_G \to G, \quad \omega = (\cdots, f_{-1}, f_0, f_1, \cdots) \mapsto f_n \circ \cdots \circ f_0.$$

As before both processes are Markov processes on the semigroup G. The proof is similar to that for the semigroup (F_A, ν).

Corollary 3.20 *Let $(X_n)_{n\in\mathbb{N}_0}$ be an aperiodic and irreducible Markov chain on a finite space A and (G, μ) a synchronising global coupling.*

i) Every constant function $g \in G_{sync}$ is an absorbing state of the backward composition process on G.

ii) $\bigcap_{n\in\mathbb{N}} \hat{F}_n(\Omega_G) = G_{sync}$ μ-almost surely.

Example 3.21 As an example let (G, μ) be the synchronising semigroup of Example 3.17.c). The semigroup generated by f_0 and f_1 is given by

$$S(f_0, f_1) := \{f_0, f_1, f_{00}, f_{11}, f_{01}, f_{10}, f_m\}$$

where

$$\begin{aligned}
f_{00} &:= f_0 \circ f_0 = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}, & f_{11} &:= f_1 \circ f_1 = \begin{bmatrix} 1 & 3 \\ 2 & 3 \\ 3 & 3 \end{bmatrix}, \\
f_{01} &:= f_0 \circ f_1 = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 2 \end{bmatrix}, & f_{10} &:= f_1 \circ f_0 = \begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 3 & 3 \end{bmatrix}, \\
f_m &:= f_1 \circ f_0 \circ f_0 = f_0 \circ f_1 \circ f_1 = \begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 3 & 2 \end{bmatrix}, \\
f_0 \circ f_1 \circ f_0 &= f_0 , & f_1 \circ f_0 \circ f_1 &= f_1 .
\end{aligned}$$

An elementary calculation shows that, indeed, this is the entire semigroup. In particular, $G_{sync} = \{ f_{00}, f_{11}, f_m \}$.

In Figure 3.3 we have illustrated backward and forward composition process on the semigroup G. In the first graph we see that constant functions are absorbing

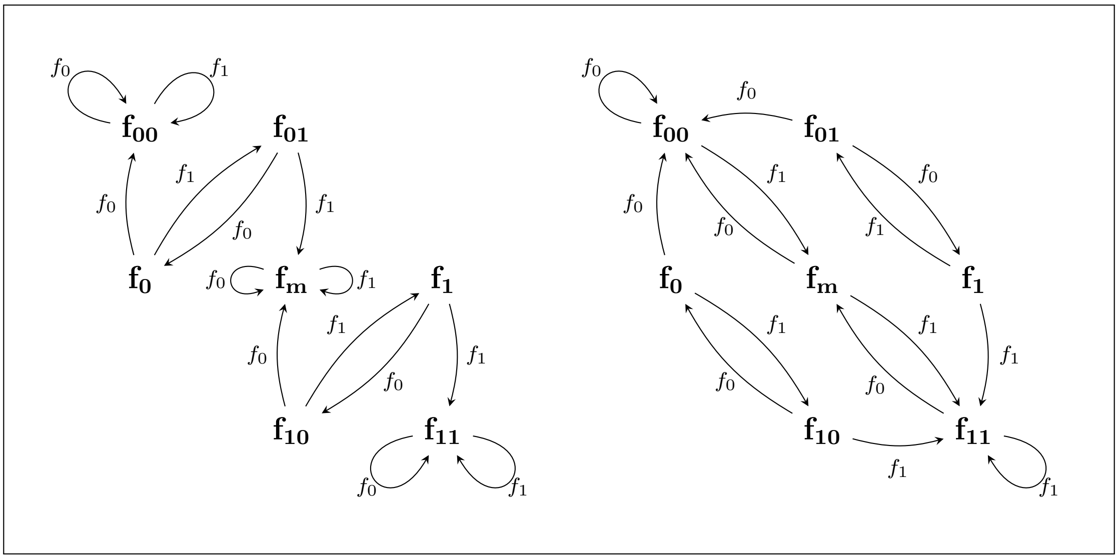


Figure 3.3: b-composition and f-composition process on (G, μ)

states. In case of the forward composition all trajectories will arrive in G_{sync} - the nodes on the main diagonal - almost surely and proceed within this subset.

By A_∞ we denote the set $A_\infty := \Big\{ \omega \in \Omega_G \colon \exists n \in \mathbb{N} \colon \hat{B}_n(\omega) \equiv \mathit{constant} \Big\}$. Although we have assigned this symbol before, there is only little chance of misinterpretation. As before the limit $\lim\limits_{n\to\infty} \hat{B}_n(\omega)$ exists for all $\omega \in A_\infty$. We refer to it by $\hat{B}_\infty(\omega)$ on A_∞.

Theorem 3.22 *Let $(X_n)_{n\in\mathbb{N}_0}$ be an aperiodic and irreducible Markov chain on a finite state space A with stationary distribution π and (G, μ) a synchronising global coupling. Then $\hat{\mu}(A_\infty) = 1$ and $\hat{B}_\infty$ is a random variable $\hat{\mu}$-almost everywhere distributed according to π.*

Proof. The proof is analogous to that of Theorem 3.10. □

Remark 3.23 One might think about the case that we are given a family of synchronising global couplings $(G_t, \mu_t)_{t\in\mathbb{N}_0}$ of a Markov chain, one for each time $t \in \mathbb{N}_0$. Even then Theorem 3.22 would be true.

3.2.1 Semigroup and Update Function

We will discuss how every synchronising global coupling yields a so called update function and vice versa. An update function is a (global) coupling of a Markov chain, here understood in terms of probability theory, with a positive probability for the trajectories to coalesce.

Definition 3.24 Let $(X_n)_{n\in\mathbb{N}_0}$ and $(Y_n)_{n\in\mathbb{N}_0}$ be two Markov chains with transition matrix T on a probability space (Ω, Σ, μ). We call a Markov chain $(Z_n)_{n\in\mathbb{N}_0} = (Z'_n, Z''_n)_{n\in\mathbb{N}_0}$ on the product space $(\Omega \times \Omega, \Sigma \otimes \Sigma, \mu_{xy})$ a *coupled Markov chain* if

i) $\mu_{xy}(Z_{n+1} \in \{ i \} \times \Omega | Z_n \in \{ j \} \times \Omega) = T_{i,j}$ and
$\mu_{xy}(Z_{n+1} \in \Omega \times \{ i \} | Z_n \in \Omega \times \{ j \}) = T_{i,j}$ for every $i, j \in A$ and all $n \in \mathbb{N}_0$;

ii) $Z'_{n_0} = Z''_{n_0}$ for some $n_0 \in \mathbb{N}_0$, then $Z'_n = Z''_n$ for all $n \geq n_0$.

The coupled Markov process Z allows us to start a Markov chain in two different states at the same time. Couplings with property ii) are sometimes called *Doeblin couplings*. For CFTP we even need to start one version of the chain in each state.

Definition 3.25 Let T be a stochastic matrix and $(X_n)_{n\in\mathbb{N}_0}$ be an aperiodic, irreducible Markov chain on a finite state space A. A mapping $\gamma : A \times C \to A$ on a probability space (C, Σ, μ) is called an *update function* if

i) $\mu(\{ c \in C \colon \gamma(i, c) = j \}) = T_{ij}$ for every $i, j \in A$,

ii) there exists an n-tupel $(c_1, \cdots, c_n)$, $c_i \in C$, and an element $a_0 \in A$, such that

$$\underbrace{\gamma(\gamma(\cdots\gamma}_{n-times}(a, c_1), \cdots), c_n) = a_0 \text{ for every } a \in A\ .$$

Assume (G, μ) is a synchronising global coupling of a Markov chain $(X_n)_{n\in\mathbb{N}_0}$ with state space A. Let $G_\mu := \{ g \in G \colon \mu(g) > 0 \}$ be the set of all elements $g \in G$ that support the measure μ. We define a mapping

$$\gamma \colon A \times G_\mu \to A\,,\ (a, g) \mapsto g(a)\ .$$

Then γ is an update function in the sense of Definition 3.25; the mapping γ conforms to both conditions of Definition 3.25:
To condition i), since

$$\mu(\{\, g \in G_\mu \colon \gamma(i,g) = j \,\}) = \mu(\{\, g \in G \colon g(i) = j \,\}) = T_{i,j} \quad \text{for every } i,j \in A$$

and to condition ii), since

$$\underbrace{\gamma(\gamma(\cdots\gamma}_{n-times}(i,g_1),\cdots),g_n) = B_n(g_1,\ldots,g_n) \equiv \mathit{constant}$$

for a finite sequence $g_1, \cdots, g_n$.
Now assume $\gamma \colon A \times C \to A$ is an arbitrary update function for the Markov chain $(X_n)_{n\in\mathbb{N}_0}$ with state space A. To each $c \in C$ there corresponds a function $f_c \in F_A$ via $C \ni c \mapsto \gamma(\cdot, c) \in F_A$. Let G_γ be the semigroup generated by the functions $\{\, f_c \colon c \in C \,\}$. The measure μ induces an image measure μ_γ on G_γ naturally. (G, μ_γ) is a global coupling, as

$$\mu_\gamma(\{\, g \in G_\gamma \colon g(i) = j \,\}) = \mu(\{\, c \in C \colon \gamma(i,c) = j \,\}) = T_{i,j}$$

and synchronising, as there exists a sequence $c_1, \ldots, c_n$, such that

$$f_{c_1} \circ \cdots \circ f_{c_n} = \underbrace{\gamma(\gamma(\cdots\gamma}_{n-times}(i,c_1),\cdots),c_n) \equiv \mathit{constant}\ .$$

As a result of this equivalence a Markov chain features a synchronising global coupling (G, μ) if and only if it features an update function γ.

Corollary 3.26 *Let $(X_n)_{n\in\mathbb{N}_0}$ be an aperiodic and irreducible Markov chain on a finite space A. Then it features an update function γ.*

3.2.2 Global Couplings and Synchronising Words

Any global coupling (G, μ) with $G \subseteq F_A$ for a finite state space A may be regarded as a road-colouring (compare Section 2.1). We begin with the graph (A, E) where the set of edges E is chosen as $E := A \times G$. The mapping $c\colon E \to G\,,\ (a,g) \mapsto g$ provides a colouring. As the colours $g \in G$ are mappings $g\colon A \to A$, this colouring is a road-colouring. If the global coupling (G,μ) is synchronising in addition, there exists a sequence of elements $(g_n,\dots,g_1)$, such that $g_1 \circ \dots \circ g_n \equiv \mathit{constant}$. This sequence translates to a synchronising word of the road-coloured graph we have constructed; for every vertex $a \in A$ the sequence $(a,g_1),(g_1(a),g_2),\dots,(g_{n-1}(\cdots g_1(a)\cdots),g_n)$ of edges targets in the same vertex $g_n(\cdots g_1(a)\cdots) = g_n \circ \dots \circ g_1(a) \in A$. Hence, if G is regarded as set of colours, a synchronising word is just a sequence of functions $(f_1,\dots,f_n)$, $f_i \in G$, whose composition is a constant function.

Proposition 3.27 *Let $(X_n)_{n\in\mathbb{N}_0}$ be an aperiodic and irreducible Markov chain.*

i) Then there exists a representation as a road-coloured graph.

ii) Then there exists a representation as a road-coloured graph with synchronising word.

Proof. Assertion i) follows immediately by the above correspondence between global coupling and road-coloured graph, and the fact that every aperiodic and irreducible Markov process has a global coupling by Proposition 3.12.
As there exists a synchronising global coupling (G,μ) by Proposition 3.16 the above correspondence yields a road-coloured graph with synchronising word. □

3.3 Résumé

Because of the correspondence of a synchronising global coupling, an update function and a c-graph with synchronising words, all assertions may be translated from one perspective to the other. In a sense to distinguish between them in the first place may seem artificial. Nevertheless, they provide different perspectives each with distinct advantages.

As for that we mention the following three equivalences:

i) CFTP terminates

if and only if

ii) there occurs a finite sequence of functions with constant composition

if and only if

iii) there occurs a synchronising word.

It is noteworthy that in the last two cases - more in iii) than in ii) - the state space itself becomes secondary in that sense that to observe coalescence of all trajectories on A, we just have to observe the sequence of colours until a synchronising word occurs. The set of colours may be regarded as set of control parameters. So for practical use it would be interesting to know all synchronising words and just have to observe the incoming sequence of colours. Unfortunately, adding an arbitrary colour to a synchronising word again yields a synchronising word. Thus, the set of all synchronising words is rather useless. Instead we would like to know a set of generating, synchronising words that are minimal. We have not found an appropriate definition of minimality, yet have some intuition on the requirements it has to meet.

CHAPTER 4

SYNCHRONISING CONVEX DECOMPOSITIONS OF OPERATORS

Composition processes and synchronising words take key positions in the perspective we have developed in Chapter 3. By establishing analogous structures on non-commutative probability spaces, we will construct a non-commutative version of the Coupling from the Past algorithm. As central component we will distinguish certain operators that we call synchronisable. We will discuss their properties, especially on finite dimensional algebras, where we are able to characterise them as aperiodic, irreducible operators. Finally, we will examine the relation between synchronisable transition operators and asymptotically complete random variables.

4.1 Synchronisable Operators

Let $\mathcal{A}$ be a C^*-algebra with identity and $T\colon \mathcal{A} \to \mathcal{A}$ be a unital, completely positive operator. The set of all unital, completely positive operators on $\mathcal{A}$ is a convex set; we denote it by $CP_u(\mathcal{A}, \mathcal{A})$ or $CP_u(\mathcal{A})$. Given a fixed finite convex decomposition $T = \sum_{i \in I} \lambda_i T_i$ with $\lambda_i \in (0,1]$ and $T_i \in CP_u(\mathcal{A}, \mathcal{A})$ for $i \in I$, we write $T_{i_{-n},\ldots,i_{-1}}$ if we refer to the composition $T_{i_{-n}} \circ \cdots \circ T_{i_{-1}}$.

Definition 4.1 We call a convex decomposition of an operator $T = \sum_{i\in I} \lambda_i T_i$ a *synchronising convex decomposition* if there exists a finite composition $T_{i_{-n},\ldots,i_{-1}}$, such that $\eta \circ T_{i_{-n},\ldots,i_{-1}} = \chi \circ T_{i_{-n},\ldots,i_{-1}}$ for every $\eta, \chi \in \mathcal{S}(\mathcal{A})$. Hence, there exists a state $\varphi \in \mathcal{S}(\mathcal{A})$, such that $\eta \circ T_{i_{-n},\ldots,i_{-1}} = \varphi$ for all $\eta \in \mathcal{S}(\mathcal{A})$. We speak of $(i_{-n}, \ldots, i_{-1})$ as a synchronising word. An operator $T\colon \mathcal{A} \to \mathcal{A}$ that allows a synchronising convex decomposition is said to be *synchronisable*.

For $\varphi \in \mathcal{S}(\mathcal{A})$ we define the conditional expectation

$$E_\varphi\colon \mathcal{A} \to \mathcal{A},\ x \mapsto \varphi(x)\mathbb{1}\ .$$

It is easy to show that the existence of a synchronising word as in the above definition is equivalent to the existence of a product $T_{i_{-n},\ldots,i_{-1}}$ of T_i, such that $T_{i_{-n},\ldots,i_{-1}} = E_\varphi$ for a $\varphi \in \mathcal{S}(\mathcal{A})$. Note that E_φ has two important properties:

Lemma 4.2 *Let $\varphi \in \mathcal{S}(\mathcal{A})$ and $S \in CP_u(\mathcal{A}, \mathcal{A})$. Then*

i) $SE_\varphi = E_\varphi$

ii) $E_\varphi S = E_{\varphi\circ S}$.

Proof. The assertions follow from $SE_\varphi(x) = \varphi(x)S(\mathbb{1}) = \varphi(x)\mathbb{1}$ and $E_\varphi S(x) = \varphi(Sx)\mathbb{1}$. □

Remark 4.3 Lemma 4.2 corresponds to Lemma 3.15. Conditional expectations E_φ, $\varphi \in \mathcal{S}(\mathcal{A})$, feature 'synchronising' properties analogous to those of the constant functions.

Proposition 4.4 *Let $S, T \in CP_u(\mathcal{A}, \mathcal{A})$ be two operators on $\mathcal{A}$ and assume T has a synchronising convex decomposition. Then $T_\lambda = \lambda T + (1-\lambda)S$, $\lambda \in (0, 1]$, has a synchronising decomposition.*

Proof. The proof is immediate. □

Remark 4.5 This result suggests that synchronisable operators exist in abundance by choosing $T = E_\varphi$ for some $\varphi \in \mathcal{S}(\mathcal{A})$.

The convex coefficients may be interpreted as measure λ on the index set I. From it we gain the product measure $\otimes_{-\mathbb{N}}\lambda$ on $\Omega := I^{-\mathbb{N}}$. For convenience we use the notation $\otimes\lambda := \otimes_{-\mathbb{N}}\lambda$ instead. To denote a finite subsequence $(\omega_{-l}, \ldots, \omega_{-k})$ for $k, l \in \mathbb{N}$, $k < l$ of an element $\omega \in \Omega$ we introduce the notation $\omega_{(-l,-k)} := (\omega_{-l}, \ldots, \omega_{-k})$.

Definition 4.6 Let $\varphi \in \mathcal{S}(\mathcal{A})$. We define the *backward n-composition*

$$B_n^\varphi \colon \Omega \to \mathcal{S}(\mathcal{A}),\ \omega \mapsto \varphi \circ T_{\omega_{(-n,-1)}} = \varphi \circ T_{\omega_{-n}} \circ \cdots \circ T_{\omega_{-1}}\ .$$

Remark 4.7 The composition $B_n \colon \Omega \to CP_u(\mathcal{A}),\ \omega \mapsto T_{\omega_{-n}} \circ \cdots \circ T_{\omega_{-1}}$ defines a forward composition on $\mathcal{A}$. With respect to a state $\varphi \in \mathcal{S}(\mathcal{A})$, however, it yields the backward composition $B_n^\varphi(\omega) = \varphi \circ B_n(\omega)$. Likewise, the composition $F_n \colon \Omega \to CP_u(\mathcal{A}),\ \omega \mapsto T_{\omega_{-1}} \circ \cdots \circ T_{\omega_{-n}}$ defines a backward composition on $\mathcal{A}$ and a forward composition on $\mathcal{S}(\mathcal{A})$. In this sense, backward and forward composition on $\mathcal{A}$ and $\mathcal{S}(\mathcal{A})$ are dually related.

We equip $\mathcal{S}(\mathcal{A})$ with the $\sigma(\mathcal{A}^*, \mathcal{A})$-topology and consider the corresponding Borel σ-algebra $\mathcal{B}(\mathcal{S}(\mathcal{A}))$ on $\mathcal{S}(\mathcal{A})$. Assume $T \colon \mathcal{A} \to \mathcal{A}$ is an operator with synchronising convex decomposition $\sum_{i \in I} \lambda_i T_i$, such that $(i_{-c}, \ldots, i_{-1})$ is a synchronising word. With respect to the product measure $\otimes\lambda$ the subset $\Omega_{sync} \subset \Omega$ of all sequences $\omega \in \Omega$ that contain the synchronising word $(i_{-c}, \ldots, i_{-1})$ is measurable and has measure $\otimes\lambda(\Omega_{sync}) = 1$. For convenience, we give a proof of this assertion: Consider the sequence of subsets $A_k := \{\, \omega \in \Omega \colon \omega_{-k-c+1} = i_{-c}, \ldots, \omega_{-k} = i_{-1} \,\}$ for $k \in \mathbb{N}$. For all $k \in \mathbb{N}$ the measure $\otimes\lambda(A_k)$ is given by $(\otimes\lambda)(A_k) = \prod_{j=1}^n \lambda_{i_j} > 0$. For this reason the sum $\sum_{k=0}^\infty \otimes\lambda(A_k)$ does not converge. By the Borel-Cantelli Lemma we conclude that the set $\limsup_{n\to\infty} A_n := \bigcap_{n=1}^\infty \bigcup_{k\geq n}^\infty A_k \subseteq \Omega_{sync}$ has measure $\otimes\lambda(\limsup_{n\to\infty} A_n) = 1$. As $\limsup_{n\to\infty} A_n \subseteq \Omega_{sync}$, the set Ω_{sync} has measure 1 as well.
On Ω_{sync} we consider the σ-algebra $\Sigma_{sync} := \{\, A \cap \Omega_{sync} \colon A \in \Sigma \,\}$. From now on we consider the backward composition B_n^φ on $\Omega_{sync} \subseteq \Omega$ and denote the restriction $B_n^\varphi|_{\Omega_{sync}}$ by B_n^φ.

Proposition 4.8 *B_n^φ is $(\Sigma_{sync}, \mathcal{B}(\mathcal{S}(\mathcal{A})))$-measurable.*

Proof. First note that the function B_n^φ attains only finitely many values in $\mathcal{S}(A)$. Therefore, only finitely many $\psi \in \mathcal{S}(\mathcal{A})$ have an inverse image unequal to the empty set $\{\emptyset\}$. These inverse images $\psi \in \mathcal{B}(S(A))$ are a union of at most finitely many cylinder sets and as such elements of Σ_{sync}. □

Lemma 4.9 *Let (X, Σ_X) and (Y, Σ_Y) be measure spaces and $X_n \in \Sigma_X$, $n \in \mathbb{N}_0$, disjoint sets, such that $X = \bigcup_n X_n$. If $f_n \colon X_n \to Y$ are $(\Sigma_X|_{X_n}, \Sigma_Y)$-measurable functions, then $f \colon X \to Y$, $f(x) := f_n(x)$ for $x \in X_n$ is (Σ_X, Σ_Y)-measurable.*

Proof. The proof is standard. □

We introduce the sets of all sequences such that the synchronising word $(i_{-c}, \dots, i_{-1})$ occurs at $-k$ for the first time:

$$P_k := \{\, \omega \in \Omega_{sync} \colon \omega_{-k-c+1} = i_{-c}, \dots, \omega_{-1} = i_{-1} \\ \text{and if } \omega_{-l-c+1} = i_{-c}, \dots, \omega_{-l} = i_{-1}\,, \text{ then } l \geq k \,\}$$

and the σ-algebras $\Sigma_{P_k} := \{\, A \cap P_k \colon A \in \Sigma \,\}$. The sets P_k are a partition of Ω_{sync}. Being $(\Sigma_{sync}, \mathcal{B}(S(A)))$-measurable, B_n^φ is $(\Sigma_{P_k}, \mathcal{B}(S(A)))$-measurable as well. Setting $(X, \Sigma_X) = (\Omega_{sync}, \Sigma_{sync})$, $(Y, \Sigma_Y) = (S(A), \mathcal{B}(S(A)))$ and $f_n = B_n^\varphi$ for a state $\varphi \in \mathcal{S}(A)$ Lemma 4.9 yields:

Corollary 4.10 *The mapping $B_\infty^\varphi \colon \Omega_{sync} \to \mathcal{S}(\mathcal{A})$, $B_\infty^\varphi(\omega) := \lim_{n\to\infty} B_n^\varphi(\omega)$ is $(\Sigma_{sync}, B(S(A)))$-measurable for every $\varphi \in \mathcal{S}(\mathcal{A})$.*

Note that, strictly speaking, the limit $\lim_{n\to\infty} B_n^\varphi(\omega)$ is artificial, because for each $\omega \in \Omega_{sync}$ there exists $n_0 \in \mathbb{N}$, such that for all $n > n_0$ $B_n^\varphi(\omega)$ is constant and does not depend on the state $\varphi \in \mathcal{S}(\mathcal{A})$.

Hence, we omit the superscript in our notation of the limit $\lim_{n\to\infty} B_n^\varphi(\omega) = B_\infty(\omega)$. For easier accessibility we cast this into a proposition:

Proposition 4.11 *For all $\varphi \in \mathcal{S}(\mathcal{A})$ the limit B_∞^φ is independent of $\varphi \in \mathcal{S}(\mathcal{A})$.*

Definition 4.12 We define θ_∞ to be the barycentre $\theta_\infty := \int_\Omega B_\infty \, d(\otimes\lambda) \in \mathcal{S}(\mathcal{A})$.

Note that $\mathcal{S}(\mathcal{A})$ is a compact convex set with respect to the $\sigma(\mathcal{A}_*, \mathcal{A})$-topology. Therefore θ_∞ is the barycentre of the measure $\otimes\lambda \circ B_\infty^{-1}$, i.e.

$$\theta_\infty(a) = \int_{\mathcal{S}(\mathcal{A})} \varphi(a)\, d(\otimes\lambda \circ B_\infty^{-1}) \ , \quad \text{for } a \in \mathcal{A} \ .$$

For information about integral representation of states we refer to [Tak03a].

Proposition 4.13 *Let $T\colon \mathcal{A} \to \mathcal{A}$ posses a synchronising convex decomposition of the form $T = \lambda E_\varphi + (1-\lambda)S$ where $\lambda \in (0,1]$, $\varphi \in \mathcal{S}(\mathcal{A})$ and $S \in CP_u(\mathcal{A})$. Then for all $\theta \in \mathcal{S}(A)$*

$$||\theta \circ T^n - \theta_\infty|| \leq 2(1-\lambda)^n \ .$$

Proof. The convex decomposition of T induces a convex combination of T^n. Applying Lemma 4.2 yields

$$\begin{aligned} T^n &= \Big(\sum_{k=1}^{n} \lambda(1-\lambda)^{k-1} E_\varphi S^{k-1}\Big) + (1-\lambda)^n S^n \\ &= \Big(\sum_{k=1}^{n} \lambda(1-\lambda)^{k-1} E_{\varphi\circ S^{k-1}}\Big) + (1-\lambda)^n S^n \ . \end{aligned}$$

Indeed, the assertion follows by an induction argument with

$$\begin{aligned} T^2 =& (\lambda E_\varphi + (1-\lambda)S)(\lambda E_\varphi + (1-\lambda)S) \\ =& \lambda^2 E_\varphi E_\varphi + (1-\lambda)\lambda S E_\varphi + \lambda(1-\lambda)E_\varphi S + (1-\lambda)^2 S^2 \\ =& \lambda^2 E_\varphi + (1-\lambda)\lambda E_\varphi + \lambda(1-\lambda)E_\varphi S + (1-\lambda)^2 S^2 \\ =& \lambda E_\varphi + \lambda(1-\lambda)E_\varphi S + (1-\lambda)^2 S^2 \end{aligned}$$

and inductive step

$$\begin{aligned} T^{n+1} =&(\lambda E_\varphi + (1-\lambda)S)T^n \\ =&\sum_{k=1}^{n}\{\lambda^2(1-\lambda)^{k-1}E_\varphi^2 S^{k-1} + \lambda(1-\lambda)(1-\lambda)^{k-1}SE_\varphi S^{k-1}\} \\ &+ \lambda(1-\lambda)^n E_\varphi S^n + (1-\lambda)^{n+1}S^{n+1} \\ =&\sum_{k=1}^{n+1}\lambda(1-\lambda)^{k-1}E_\varphi S^{k-1} + \lambda(1-\lambda)^n E_\varphi S^n + (1-\lambda)^{n+1}S^{n+1} \\ =&\sum_{k=1}^{n+1}\lambda(1-\lambda)^{k-1}E_\varphi S^{k-1} + (1-\lambda)^{n+1}S^{n+1} \ . \end{aligned}$$

Set $T_1 = E_\varphi$, $T_2 = S$ and $\lambda_1 = \lambda$. Then in the above notation $\Omega = \{1,2\}^{-\mathbb{N}}$. The cylinder sets P_k are given by

$$P_k := \{\omega \in \Omega\colon T_{\omega_{-k}} = E_\varphi,\ T_{\omega_{-l}} = S,\ l = k-1,\dots,1\}$$

For $k = 1,\dots,n$ they correspond one to one to a single summand of T^n containing E_φ.

For every positive element $a \in \mathcal{A}_+$ the sequence $\left(\chi_{P_k}(\omega)B_\infty(\omega)(a)\right)_{k\in\mathbb{N}}$, where χ_{P_k} is the characteristic function on P_k, is a sequence of non-negative functions that converge to $B_\infty(a)$ pointwisely. From the Monotone Convergence Theorem follows

$$\theta_\infty(\omega) = \int\limits_{\Omega_{sync}} B_\infty(\omega)\, d(\otimes\lambda) = \lim_{n\to\infty}\sum_{k=1}^{n}\int\limits_{P_k} B_\infty(\omega)\, d(\otimes\lambda) \ .$$

Evaluating B_∞ on sets P_k yields

$$\lim_{n\to\infty}\sum_{k=1}^{n}\int\limits_{P_k} B_\infty(\omega)\, d(\otimes\lambda) = \lim_{n\to\infty}\sum_{k=1}^{n}\lambda(1-\lambda)^{k-1}E_\varphi S^{k-1} \ .$$

Therefore, we can estimate the difference in norm for $\theta \in \mathcal{S}(\mathcal{A})$ by

$$
\begin{aligned}
||\theta \circ T^n - \theta_\infty|| = ||\theta \circ \left(\sum_{k=1}^{n} \lambda(1-\lambda)^{k-1} E_\varphi S^{k-1} \right) + (1-\lambda)^n \theta \circ S^n - \theta_\infty|| \\
=& ||\theta \circ \sum_{k=1}^{n} \lambda(1-\lambda)^{k-1} E_\varphi S^{k-1} + (1-\lambda)^n \theta \circ S^n \\
& - \theta \circ \sum_{k=1}^{\infty} \lambda(1-\lambda)^{k-1} E_\varphi S^{k-1}|| \\
\leq& ||(1-\lambda)^n \theta \circ S^n|| + || \sum_{k=1}^{n} \lambda(1-\lambda)^{k-1} E_\varphi S^{k-1} \\
& \qquad - \sum_{k=1}^{\infty} \lambda(1-\lambda)^{k-1} E_\varphi S^{k-1}|| \\
=& ||(1-\lambda)^n \theta \circ S^n|| + || \sum_{k=n+1}^{\infty} \lambda(1-\lambda)^{k-1} E_\varphi S^{k-1}|| \\
\leq& (1-\lambda)^n + \lambda(1-\lambda)^n \sum_{k=0}^{\infty} (1-\lambda)^k \leq 2(1-\lambda)^n \ .
\end{aligned}
$$

□

Remark 4.14 In general, the above Proposition does not apply to an operator $T\colon \mathcal{A} \to \mathcal{A}$ with a synchronising convex decomposition. However, if $T = \sum_{i \in I} \lambda_i T_i$ with a synchronising word $(i_{-c}, \ldots, i_{-1})$, then the c-th power of T is given by $T^c = \lambda_\pi E_\varphi + (1-\lambda_\pi) S$ for a state $\varphi \in \mathcal{S}(\mathcal{A})$, an operator $S \in CP_u(\mathcal{A}, \mathcal{A})$ and $\lambda_\pi = \prod_{k=1}^{c} \lambda_{i_{-k}}$. Proposition 4.13 applies to T^c:

$$||\theta \circ (T^c)^m - \theta_\infty|| \leq 2(1-\lambda_\pi)^m.$$

Definition 4.15 A state $\theta \in \mathcal{S}(\mathcal{A})$ is called *absorbing* if for all $\varphi \in \mathcal{S}(\mathcal{A})$ and $a \in \mathcal{A}$

$$\lim_{n\to\infty} \varphi \circ T^n(a) = \theta(a) \ .$$

Corollary 4.16 *Let $T\colon \mathcal{A} \to \mathcal{A}$ be a unital, completely positive operator with synchronising convex decomposition. Then the barycentre $\theta_\infty \in \mathcal{S}(\mathcal{A})$ is an absorbing state. In particular, θ_∞ is a unique invariant state of T.*

If $T\colon \mathcal{A} \to \mathcal{A}$ is a synchronsiable operator with a pure invariant state $\theta \in \mathcal{S}(\mathcal{A})$, the state θ cannot be a true barycentre.

Proposition 4.17 *Let $T\colon \mathcal{A} \to \mathcal{A}$ posses a synchronising convex decomposition and θ_∞ be a pure state. If $T = \sum_{i\in I} \lambda_i T_i$ is an arbitrary convex decomposition, then $\theta_\infty \circ T_i = \theta_\infty$ for all $i \in I$.*

Proof. Let θ_∞ be a pure state. Suppose there exists one term T_k such that $\theta_\infty \circ T_k = \theta_k \neq \theta_\infty$. Then we can obtain a convex decomposition of θ_∞ according to the following line

$$\theta_\infty = \theta_\infty \circ T = \sum_{i\neq k} \lambda_i \theta_\infty \circ T_i + \lambda_k \theta_\infty \circ T_k = (1-\lambda_k)\theta_\infty + \lambda_k \theta_k \ .$$

This, however, leads to a contradiction. □

Corollary 4.18 *Let $T\colon \mathcal{A} \to \mathcal{A}$ posses a synchronising convex decomposition and $\theta_\infty \in \mathcal{S}(\mathcal{A})$ be a pure state. If $N \in \mathbb{N}$ is a natural number, such that $T^N = \lambda E_\varphi + (1-\lambda)S$ for some state $\varphi \in \mathcal{S}(\mathcal{A})$ and $S \in CP_u(\mathcal{A},\mathcal{A})$, then $\varphi = \theta_\infty$ and $\theta_\infty \circ S = S$.*

Proposition 4.19 *Let $\mathcal{A}$ be a W^*-algebra and $T\colon \mathcal{A} \to \mathcal{A}$ a normal, unital and completely positive operator with synchronising convex decomposition. Then the invariant state θ_∞ is normal.*

Proof. For every state $\theta \in \mathcal{S}(\mathcal{A})$ the sequence of states $(\theta \circ T^n)_{n\in\mathbb{N}}$ converges to θ_∞ uniformly by Proposition 4.13. If θ is a normal state, $(\theta \circ T^n)_{n\in\mathbb{N}}$ is a sequence of normal states. As the set of normal states is uniformly closed, its uniform limit θ_∞ is normal, too. □

We gather all our findings in the next theorem.

Theorem 4.20 *Let $T\colon \mathcal{A} \to \mathcal{A}$ be a synchronisable operator with synchronising convex decomposition $T = \sum_{i\in I} \lambda_i T_i$ on a C^*-algebra with identity. Then:*

i) The operator T possesses a unique invariant state.

ii) The invariant state of T and the barycentre

$$\theta_\infty = \int_{\mathcal{S}(\mathcal{A})} \varphi(a)\, d(\otimes\lambda \circ B_\infty^{-1})$$

coincide.

iii) The invariant state θ_∞ is absorbing, and the rate of convergence is exponential.

iv) If additionally $\mathcal{A}$ is a W^*-algebra and T is a normal operator, then θ_∞ is a normal state.

Proof. If $T\colon \mathcal{A} \to \mathcal{A}$ possesses a synchronising convex decomposition of the form $T = \lambda E_\varphi + (1-\lambda)S$, assertion iii) follows from Proposition 4.13. In general, if $T = \sum_{i\in I} \lambda_i T_i$ with a synchronising word $(i_{-c}, \ldots, i_{-1})$, we have the following inequality by Remark 4.14

$$||\theta \circ (T^c)^m - \theta_\infty|| \leq 2(1-\lambda_\pi)^m \ .$$

From this inequality we obtain the following estimate for the operator T

$$\begin{aligned} ||\theta \circ T^m - \theta_\infty|| =& ||\theta \circ T^{\lfloor \frac{m}{c} \rfloor c + r} - \theta_\infty|| = ||(\theta \circ T^{\lfloor \frac{m}{c} \rfloor c} - \theta_\infty) \circ T^r|| \\ \leq& ||(\theta \circ T^{\lfloor \frac{m}{c} \rfloor c} - \theta_\infty|| \leq 2(1-\lambda_\pi)^{\lfloor \frac{m}{c} \rfloor} \ , \end{aligned}$$

where $\lfloor\ \rfloor\colon \mathbb{R} \to \mathbb{N}_0,\ x \mapsto \max\{\, n \in \mathbb{N}_0\colon n \leq x \,\}$ is the floor function. □

4.2 Synchronisable Operators on M_n

For a classical Markov transition matrix it is sufficient to ask for aperiodicity and irreducibility to assure the existence of a synchronising convex decomposition. We will see that any unital, irreducible and completely positive operator on a finite dimensional von Neumann algebra $\mathcal{A}$ allows a synchronising convex decomposition as well, as long as its invariant state is absorbing. By Lemma 2.23, this is the case if the invariant state is pure. According to Proposition 4.17 a pure invariant state enforces very restricting conditions on any convex decomposition of a synchronisable operator. Excluding those operators leaves us with the question which of them possess a faithful absorbing state. Recall from Section 2.2.2 that in classical probability theory an irreducible transition matrix T has a faithful 'absorbing' measure if and only if T is aperiodic. Similarly, in non-commutative probability theory this requirement is met for an irreducible operator if and only if it is aperiodic in the sense of F. Fagnola and R. P. Bidot [FB09]. Therefore, our version of the non-commutative Coupling from the Past algorithm reproduces the commutative situation on finite dimensional von Neumann algebras nicely.

4.2.1 Arveson's Theorem

Theorem 4.21 *[Sti55] Let $\mathcal{A}$ be a finite algebra and $T\colon \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a completely positive operator. Then there exist a representation $(\pi, \mathcal{K})$ of $\mathcal{A}$ on a Hilbert space $\mathcal{K}$ and an operator $V\colon \mathcal{H} \to \mathcal{K}$ such that*

$$T(a) = V^*\pi(a)V \ .$$

We call the triple $(\mathcal{K}, \pi, V)$ the *Stinespring representation* of the operator T. If $\{\, \pi(a)V\xi\colon a \in \mathcal{A},\ \xi \in \mathcal{H} \,\}$ is dense in $\mathcal{K}$, the Stinespring representation $(\mathcal{K}, \pi, V)$ is called *cyclic*. Any two cyclic Stinespring representations $(\mathcal{K}_1, \pi_1, V_1)$ and $(\mathcal{K}_2, \pi_2, V_2)$ of T are unitarily equivalent. In particular, $\dim \mathcal{K}_1 = \dim \mathcal{K}_2$.
For $T \in CP(\mathcal{B}(\mathcal{H}), \mathcal{B}(\mathcal{H}))$ the completely positive operators on $\mathcal{B}(\mathcal{H})$, the set $\{\, S \in CP(\mathcal{B}(\mathcal{H}))\colon 0 \leq S \leq T \,\}$ is convex. We denote it by $[0, T]$. Given $T\colon \mathcal{A} \to \mathcal{B}(\mathcal{H})$ with Stinespring representation $\{\mathcal{K}, \pi, V\}$, for each $a' \in \pi(\mathcal{A})'$ we define the operator

$$V_{a'}\colon \mathcal{A} \to \mathcal{B}(\mathcal{H}),\ a \mapsto V^* a' \pi(a) V \ .$$

Theorem 4.22 *[Arv69] The mapping $a' \mapsto V_{a'}$ is an affine order isomorphism of the partially ordered convex set of operators $\{a' \in \pi(\mathcal{A})'\colon 0 \leq a' \leq \mathbb{1}\}$ onto $[0, T]$.*

For the rest of this section we will restrict ourselves to $\mathcal{A} = M_n$. Let $T\colon M_n \to M_n$ be a completely positive operator and consider M_n as space of operators $\mathcal{B}(\mathcal{H})$ on $\mathcal{H} = \mathbb{C}^n$. Given any representation $\pi\colon M_n \to M_k$ of M_n, there exists a natural number $m \in \mathbb{N}$, such that $k = nm$ and $M_k \simeq M_n(M_m)$. Hence, we consider M_n represented on $\mathcal{K} = \mathcal{H} \otimes \mathbb{C}^m$ and choose the following representation of $\mathcal{A}$ on $\mathcal{K}$

$$\pi(x) = x \otimes \mathbb{1} \ .$$

We may choose the following identification:

$$\mathcal{K} = \oplus_{k=1}^m \mathcal{H} \quad \text{and} \quad \pi(x) = \oplus_{k=1}^m x \ .$$

Now, if T is represented by a Stinespring representation $(\mathcal{K}, \pi, V)$, the operator V may be written

$$V\colon \mathcal{H} \to \oplus_{i=1}^m \mathcal{H},\ \xi \mapsto (a_1\xi, \dots, a_m\xi) \ .$$

Consequently

$$T(a) = V^*\pi(x)V = (a_1^*, \dots, a_n^*) \begin{pmatrix} x & & \\ & \ddots & \\ & & x \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \sum_{i=1}^m a_i^* x a_i \ .$$

This representation of a completely positive operator T is often called *Kraus decomposition* of T. We will refer to it as *concrete representation*. If there is no chance for misinterpretation, we call the set $\{ a_i \}_{i \in I}$ a concrete representation of a completely positive operator T as well. Given another completely positive operator $S \in [0, T]$ there exists an element $a' \in \pi(M_n)'$, such that $S(x) = V^* a'(x \otimes \mathbb{1})V$. The element a' is sometimes called Radon-Nikodym density. As $\pi(M_n)' \simeq \mathbb{1} \otimes M_m$ the element a' may be written as $a' = \mathbb{1} \otimes \Lambda * \Lambda$ for some $\Lambda = (\lambda_{ij})_{i,j} \in M_m$. Since

$\mathbb{1} \otimes \Lambda^* \in \pi(M_n)'$ and $\mathbb{1} \otimes \Lambda \in \pi(M_n)'$ we obtain

$$\begin{aligned} S(x) =& V^*(\mathbb{1} \otimes \Lambda)^* \pi(x)(\mathbb{1} \otimes \Lambda)V = V^*(\mathbb{1} \otimes \Lambda)^* \pi(x)(\lambda_{ij}\mathbb{1}_{M_n})_{i,j} \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \\ =& \begin{pmatrix} \sum_{i=1}^m \lambda_{1i} a_i \\ \vdots \\ \sum_{i=1}^m \lambda_{mi} a_i \end{pmatrix}^* \pi(x) \begin{pmatrix} \sum_{i=1}^m \lambda_{1i} a_i \\ \vdots \\ \sum_{i=1}^m \lambda_{mi} a_i \end{pmatrix} . \end{aligned}$$

Thus, we see that the concrete representation of S may be obtained as a linear combination of $a_i \in M_m$:

$$S(x) = \sum_{i=1}^{m} b_i^* x b_i \quad \text{with } b_i = \sum_{j=1}^{m} \lambda_{ij} a_j \ .$$

4.2.2 Full Concrete Representation

Definition 4.23 Let $T\colon M_n \to M_n$ be a completely positive operator and $T(x) = \sum_{i \in I} a_i^* x a_i$ a concrete representation. If the set of $\{ a_i \}_{i \in I}$ spans the entire space linearly, i.e. $\operatorname{span}\{ a_i \colon i \in I \} = M_n$, the representation is said to be *full*. Featuring a full concrete representation, we call an operator T *full* likewise.

Proposition 4.24 *Let $T\colon M_n \to M_n,\ T(x) = \sum_{i \in I} a_i^* x a_i$ be a (unital) completely positive operator. The concrete representation is full if and only if for every (unital) completely positive operator $S\colon M_n \to M_n$ there exists a real number $\alpha \in \mathbb{R}$ ($\alpha \in (0,1]$), such that $\alpha S \leq T$.*

Proof. Let the operator S be represented by

$$S(x) = \sum_{i=1}^{k} b_i^* x b_i, \quad \text{for a } k \leq n^2 \ .$$

Since the operator T is full, we can represent the elements b_i as linear combinations of the elements a_i. By reversing the procedure described at the end of Section 4.2.1, we can determine the density $\Lambda^*\Lambda$:

$$S = V^*\Lambda^*\Lambda\pi(x)V \ .$$

By choosing a real number $\alpha \in \mathbb{R}$, such that $\alpha\Lambda^*\Lambda \leq \mathbb{1}$ we make certain $\alpha S \leq T$:

$$\alpha S(x) = V^*(\alpha\Lambda^*\Lambda)\pi(x)V \leq V^*\pi(x)V = T(x) \quad \text{for all } x \in (M_n)_+ \ .$$

Assume for every completely positive operator $S\colon M_n \to M_n$ there exists a real number $\alpha > 0$, such that $\alpha S \leq T$. Given any operator αS by the mentioned procedure (cf. section 4.2.1) we obtain a concrete representation $\{ b_j \}_{j\in J}$ of αS with $b_j = \sum_{i\in I} \lambda_{ij} a_i$. □

Corollary 4.25 *Let $T\colon \mathcal{A} \to \mathcal{A}$ be a unital, completely positive and full operator. Then for each $\varphi \in \mathcal{S}(\mathcal{A})$ there exists a synchronising convex decomposition of the operator T of the form*

$$T = \lambda E_\varphi + (1-\lambda)R_\varphi \ , \ \lambda \in (0,1] \, ,$$

such that $R_\varphi(\mathbb{1}) = \mathbb{1}$.

Proof. By Proposition 4.24 we know there exists a real number $\lambda \in (0,1]$, such that $T \geq \lambda E_\varphi$. Therefore, we can decompose T into

$$T = \lambda E_\varphi + (T - \lambda E_\varphi) \ .$$

In particular, $T - \lambda E_\varphi \geq 0$. We set $R_\varphi = \frac{1}{1-\lambda}(T - \lambda E_\varphi)$. The operator R_φ is unital and

$$T = \lambda E_\varphi + (1-\lambda)R_\varphi \ .$$

□

This is a very valuable piece of information. It provides us firstly with a direct way to tell if an operator allows a synchronising convex decomposition and secondly with a method to acquire synchronising decompositions. Albeit, a lot of operators do not have a full concrete representation, there exist many operators whose powers do. Being mainly interested in asymptotic behaviour, it is not a loss to study the latter instead.

Proposition 4.26 *Let $T\colon M_n \to M_n$ be a unital, completely positive operator and $T(x) = \sum_{i\in I} a_i^* x a_i$ a concrete decomposition. Then the following assertions are equivalent.*

a) The operator T is full.

b) For every state $\varphi \in \mathcal{S}(M_n)$ the state $\varphi \circ T$ is faithful.

Proof. **a) ⇒ b)** By Proposition 4.24 for every $S \in CP_u(M_n, M_n)$ there exists a real number $\alpha > 0$, such that $T \geq \alpha S$. Hence, for every state $\varphi \in \mathcal{S}(M_n)$ we obtain

$$\varphi \circ (T - \alpha S)(x) \geq 0\ , \text{ for all } x \in (M_n)_+.$$

Given an element $x \in (M_n)_+$ there exists a state $\theta \in \mathcal{S}(M_n)$ with support projection $p_\theta := \operatorname{supp}\theta$, such that $p_\theta x p_\theta \neq 0$. This yields $\varphi \circ T(x) \geq \alpha\varphi \circ E_\theta(x) > 0$ for some $\alpha \in \mathbb{R}^+$ with $\alpha E_\theta \leq T$.

b) ⇒ a) Let the set L be given by $L := \{\, x \in M_n \colon ||x|| = 1, x \geq 0 \,\}$ and the set K by $K := \mathcal{S}(M_n) \times L$. This set is compact with respect to the product topology. The mapping $\mathcal{S}(M_n) \times L \ni (\varphi, x) \mapsto \varphi \circ T(x)$ is continuous and strictly positive, i.e. $\varphi \circ T(x) > 0$ for every $x \in L$, as $\varphi \circ T$ is faithful. Hence, the minimum $\alpha := \min_{(\varphi,x)\in K)} \{\, \varphi \circ T(x) \,\}$ exists and is strictly positive.
For all $x \in L$ follows

$$\frac{1}{\alpha}\varphi \circ T(x) \geq 1 \geq \varphi \circ S(x)\ ,$$

and therefore

$$\varphi \circ (T - \alpha S)(x) \geq 0\ .$$

With this inequality being true for every $\varphi \in \mathcal{S}(M_n)$, it follows that $(T - \alpha S)(x) \geq 0$ for all $x \in L$; or equivalently $T \geq \alpha S$. $\square$

Definition 4.27 Let $\{a_i\}_{i\in I}$ be a concrete representation of a completely positive, unital operator T, such that $\{a_i\}_{i\in I}$ is linearly independent. Then we define the *cs-rank* to be

$$\operatorname{rank}_{cs} T = |I| \ .$$

Remark 4.28 The abbreviation *cs* stands for concrete and Stinespring. As any cyclic Stinespring representation of an operator T is cyclic if and only if the associated concrete representation is linearly independent, and as cyclic Stinespring representations of T are unitarily equivalent, the cs-rank does not depend on the concrete representation. The cs-rank is maximal if and only if the operator is full; in that case $\operatorname{rank}_{cs} T = n^2$.

Example 4.29 Let $\varphi \in \mathcal{S}(M_n)$ be a pure state. Then the conditional expectation

$$M_n \ni x \mapsto \varphi(x)\mathbb{1} \in M_n$$

has $\operatorname{rank}_{cs} E_\varphi = n$.

Proof. Since φ is a pure state, there exists a vector $\zeta \in \mathbb{C}^n$, such that the density matrix ρ_φ of the state φ may be written as $\rho = (\zeta_i\overline{\zeta_j})_{ij}$. Then the set $\{a_i \colon i = 1, \ldots, n\}$ with $a_i = \sum_{i=1}^n \overline{\zeta}_i E_{ij}$ gives a concrete representation of E_φ, where E_{ij} denotes the matrix unit $E_{ij} := (\delta_{i,k}\delta_{j,l})_{l,k} \in M_n$. Obviously, this set is linearly independent. □

Example 4.30 Let $\varphi \in \mathcal{S}(\mathcal{A})$ be a faithful state. Then the conditional expectation

$$M_n \ni x \mapsto \varphi(x)\mathbb{1} \in M_n$$

has $\operatorname{rank}_{cs} E_\varphi = n^2$.

Lemma 4.31 *[Cho75] Let $T \colon M_n \to M_n$ be a completely positive operator. Then the cs-rank of T is equal to the matrix rank of $\Phi(T) := (T(E_{ij}))_{ij}$*

$$\operatorname{rank}_{cs} T = \operatorname{rank} \Phi(T) \ .$$

In particular, the operator T possesses a full concrete representation if and only if the matrix $\Phi(T)$ has full matrix rank.

Remark 4.32 The matrix $\Phi(T)$ is well known under the name of Choi matrix [Cho75].

Proof. The following proof is kept in a rather sketchy manner. Although it is not proposed in [Cho75] explicitly, all necessary details to prove this assertion may be found there.
Let $\{a_i\}_{i\in I}$ be a concrete representation of the operator T, such that the set $\{\, a_i \,\}_{i\in I}$ is linearly independent. The matrix $\Phi(T)$ may be written as $\Phi(T) = \sum_{i\in I} \zeta_i^* \zeta_i$ for vectors $\zeta_i^* \in (\mathbb{C}^n)^n$. The $n \times 1$-blocks the vectors ζ_i^* may be thought of being divided into are given by the columns of the respective matrices a_i:

$$a_i^* = \left(\zeta_{i1}^*, \ldots, \zeta_{in}^*\right) \ .$$

The vectors $\{\, \zeta_i^* \,\}_{i\in I}$ are linearly independent only if the matrices $\{a_i\}_{i\in I}$ are linearly independent. This, of course, is true for the matrices $\{\zeta_i^* \zeta_i\}_{i\in I}$ as well. □

Proposition 4.33 *Let $T\colon M_m \to M_m$ be a unital, completely positive operator, such that $\lim_{n\to\infty} T^n = E_\varphi$ for a faithful state $\varphi \in \mathcal{S}(M_m)$. Then there exists a natural number $n_0 \in \mathbb{N}$, such that*

$$\operatorname{rank}_{cs} T^n = m^2 \quad \textit{for all } n \geq n_0 \ .$$

Proof. The matrix $\Phi(E_\varphi)$ is invertible. Therefore $\Phi(E_\varphi)$ is an element of the open set of invertible matrices GL_{m^2}. Since $\lim_{n\to\infty} \Phi(T^n) = \Phi(E_\varphi)$ with respect to any topology induced by a matrix norm, there exists a natural number $n_0 \in \mathbb{N}$, such that $\Phi(T^n) \in GL_{m^2}$ for all $n \geq n_0$. □

Theorem 4.34 *Let $T\colon M_n \to M_n$ be an irreducible operator with an absorbing state $\varphi \in \mathcal{S}(M_n)$. Then there exists a natural number $n_0 \in \mathbb{N}$, such that T^{n_0} possesses a synchronising convex decomposition.*

Proof. The sequence T^n converges to the conditional expectation E_φ. Since φ is faithful (Lemma 2.21), the operator E_φ is full. Therefore, there exists a natural number n_0, such that T^n is full for all $n \geq n_0$. Being a full operator T^{n_0} possesses a synchronising convex decomposition. $\square$

Corollary 4.35 *Let $T\colon M_n \to M_n$ be an aperiodic, irreducible operator. Then there exists a natural number $n_0 \in \mathbb{N}$, such that T^{n_0} possesses a synchronising convex decomposition.*

This corollary stresses another analogy to the commutative situation: To every transition matrix M of an aperiodic, irreducible Markov chain there may be found a natural number n, such that the n-th power M^n of M has strictly positive entries.

Proposition 4.36 *Let $T\colon M_n \to M_n$ be an aperiodic, irreducible operator. Then there exists a natural number $n \in \mathbb{N}$, such that for every state $\varphi \in \mathcal{S}(M_n)$ the state $\varphi \circ T^n$ is faithful.*

Proof. By Corollary 4.35 there exists a natural number $n \in \mathbb{N}$, such that T^n is full. Therefrom the assertion follows by Proposition 4.26. $\square$

4.3 Synchronisable Operators on Finite Algebras

The results we have presented in the case $\mathcal{A} = M_n$ may be applied to prove corresponding statements for finite algebras. By the well-known identification of a finite algebra $\mathcal{A}$ with a finite direct sum $\bigoplus_k M_{n_k}$, we regard $\mathcal{A}$ as a subalgebra of some larger matrix algebra M_n, where $n = \sum_k n_k$. The conditional expectation E_A onto the subalgebra $\mathcal{A} \subseteq M_n$ is uniquely determined and faithful, i.e. $E_A(x) > 0$ for every $x \in (M_n)_+$. Given an operator $T\colon \mathcal{A} \to \mathcal{A}$ the composition with the conditional expectation $E_\mathcal{A}$ is a natural extension $\hat{T}\colon M_n \to M_n$, $\hat{T} := T \circ E_\mathcal{A}$.

Definition 4.37 A unital, completely positive operator $T\colon \mathcal{A} \to \mathcal{A}$ on a finite algebra $\mathcal{A}$ is called full if $\varphi \circ T$ is faithful for every $\varphi \in \mathcal{S}(\mathcal{A})$.

Definition 4.38 If $T\colon \mathcal{A} \to \mathcal{A}$ is an operator on $\mathcal{A} \subseteq M_n$ we define the operator $\hat{T}\colon M_n \to M_n$ by

$$\hat{T}(x) := T \circ E_A(x) \quad \text{for } x \in M_n \ .$$

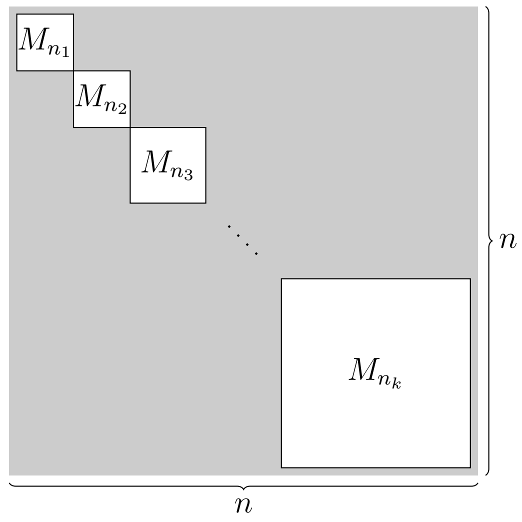


Figure 4.1: $\mathcal{A} = \bigoplus_k M_{n_k}$

Theorem 4.39 *Let $T\colon \mathcal{A} \to \mathcal{A}$ be a unital, completely positive operator on $\mathcal{A}$. Then the following assertions are equivalent:*

a) The operator $T\colon \mathcal{A} \to \mathcal{A}$ is full.

b) The operator $\hat{T}\colon M_n \to M_n$ is full.

Proof. **a) $\Rightarrow$ b)** Since the conditional expectation E_A is faithful, $E_A(x) > 0$ for every element $x \in (M_n)_+$. With T being a full operator, $\varphi \circ T(E_A(x)) > 0$ for every state $\varphi \in \mathcal{S}(M_n)$ and $x \in (M_n)_+$. It follows that

$$(\varphi \circ \hat{T})(x) = \varphi \circ T(E_A(x)) > 0 \ .$$

b) $\Rightarrow$ a) Assume $\hat{T}$ is a full operator. Then for every $\varphi \in \mathcal{S}(\mathcal{A})$ and $x \in \mathcal{A}$ we obtain

$$\varphi \circ T(x) = (\varphi \circ E_A) \circ T \circ E_A(x) = (\varphi \circ E_A) \circ \hat{T}(x) \ .$$

Since $(\varphi \circ E_A) \circ \hat{T}$ is faithful, $\varphi \circ T$ is faithful. $\square$

Theorem 4.40 *Let $T\colon \mathcal{A} \to \mathcal{A}$ be a unital, completely positive operator on $\mathcal{A}$. Then the following assertions are equivalent:*

a) The operator $T\colon \mathcal{A} \to \mathcal{A}$ is aperiodic and irreducible.

b) The operator $\hat{T}\colon M_n \to M_n$ is aperiodic and irreducible.

Proof. **a) $\Rightarrow$ b)** Assume $\hat{T}(x) = \lambda x$ for some element $x \in \mathcal{A}$ and $\lambda \in \mathbb{C}$. As λx is an element of the image $\operatorname{im} T \subseteq \mathcal{A}$ of T, x is an element of $\mathcal{A}$. Consequently, $T(x) = \hat{T}(x) = \lambda x$. By Theorem 2.32 the fixed space of the operator T is $\mathcal{F}(T) = \mathbb{C}\mathbb{1}$ and $\sigma(T) \cap \mathbb{T} = \{1\}$. Therefore, the fixed space of $\hat{T}$ is $\mathcal{F}(\hat{T}) = \mathcal{F}(T) = \mathbb{C}\mathbb{1}$ and $\sigma(\hat{T}) \cap \mathbb{T} = \sigma(T) \cap \mathbb{T} = \{1\}$, too.

b) $\Rightarrow$ a) The proof is immediate. □

Proposition 4.41 *Let $T\colon \mathcal{A} \to \mathcal{A}$ be a unital, completely positive operator, such that $\lim_{n\to\infty} T^n = E_\varphi$ for a faithful state $\varphi \in \mathcal{S}(\mathcal{A})$. Then there exists a natural number $n_0 \in \mathbb{N}$, such that T^n is full for all $n \geq n_0$.*

Proof. Instead of the operator $T\colon \mathcal{A} \to \mathcal{A}$ we consider the operator $\hat{T}\colon M_n \to M_n$. The powers $\hat{T}^k$ converge to E_φ as well: $\lim\limits_{k\to\infty} \hat{T}^k = \lim\limits_{k\to\infty} (T \circ E_A)^k = \lim\limits_{k\to\infty} T^k \circ E_A = E_\varphi$. As $\hat{T}$ is an operator on a matrix algebra we can apply Lemma 4.33. □

Theorem 4.42 *Let $T\colon \mathcal{A} \to \mathcal{A}$ be an irreducible operator with an absorbing state $\varphi \in \mathcal{S}(\mathcal{A})$. Then there exists a natural number $n_0 \in \mathbb{N}$, such that T possesses a synchronising convex decomposition.*

Proof. Again, we proof the assertion for the operator $\hat{T}\colon M_n \to M_n$ in place of T. By Proposition 4.41 and Theorem 4.39 it follows that there exists a natural number n_0, such that $\hat{T}^n$ is full for all $n \geq n_0$. We conclude that the operator $\hat{T}$ possesses a synchronising convex decomposition. □

4.4 Asymptotic Completeness and Synchronisation

As we have seen in Section 2.3.2 an asymptotically complete random variable $J\colon (\mathcal{A},\varphi) \to (\mathcal{A}\otimes\mathcal{C},\varphi\otimes\psi)$ yields an irreducible transition operator $T_\psi : (\mathcal{A},\varphi) \to (\mathcal{A},\varphi)$ that has an absorbing state. In the light of this, the question raises up if an asymptotically complete random variable has a synchronisable transition operator.

Theorem 4.43 *Let $(\mathcal{A},\varphi)$ be a finite dimensional probability space and $J\colon (\mathcal{A},\varphi) \to (\mathcal{A}\otimes\mathcal{C},\varphi\otimes\psi)$ an asymptotically complete random variable. Then there exists a natural number $n_0 \in \mathbb{N}$, such that $T_\psi^{n_0}$ is synchronisable.*

Proof. According to Proposition 2.30 the transition operator T_ψ possesses an absorbing state $\varphi \in \mathcal{S}(\mathcal{A})$. It then follows by Theorem 4.42 that $T_\psi^{n_0}$ has a synchronising convex decomposition for some natural number $n_0 \in \mathbb{N}$. □

As before we now consider a unital, completely positive operator $T\colon \mathcal{A} \to \mathcal{A}$ with a fixed synchronising convex decomposition $T = \sum_{i\in I} \lambda_i T_i$. We choose $\varphi = \theta_\infty$, the barycentre defined in Definition 4.12, and put $\mathcal{C} := L^\infty(I,\lambda)$, where λ is the probability measure on I given by $\lambda(i) := \lambda_i$ and

$$\mathcal{J}\colon (\mathcal{A},\theta_\infty) \to (\mathcal{A}\otimes\mathcal{C},\theta_\infty\otimes\psi_\lambda),\ a \mapsto \sum_i T_i(a)\otimes\varepsilon_i\,,$$

where $\psi_\lambda(c) := \int_I c\,d\lambda,\ c \in \mathcal{C}$, and $\{\,\varepsilon_i\colon i = 1,\dots,|I|\,\}$ is a canonical vector base of $\mathcal{C} \simeq \mathbb{C}^{|I|}$. Be aware that we have dropped the prerequisite of $\mathcal{J}$ to be a random variable, i.e. a *-homomorphism. Now, we construct A^+, $\mathcal{J}^+$, φ^+ and i^+. We do so in analogy to the coupling construction in section 1.2.10:

$$\mathcal{A}^+ := \mathcal{A}\otimes\mathcal{C}^+ := \mathcal{A}\otimes\bigotimes_{\mathbb{N}_0}\mathcal{C}, \quad \varphi^+ = \varphi\otimes(\otimes_{\mathbb{N}_0}\psi),$$
$$i^+\colon \mathcal{A}\to\mathcal{A}^+,\ a\mapsto a\otimes\mathbb{1}, \qquad \mathcal{J}^+\colon \mathcal{A}^+\to\mathcal{A}^+,\ a\otimes c\mapsto \mathcal{J}(a)\otimes c\,.$$

Remark 4.44 As we have mentioned before, in general, $\mathcal{J}$ is not a random variable. So the definition of asymptotic completeness should not be used for $\mathcal{J}$.

Yet, the definition as we have stated it does not necessarily require $\mathcal{J}$ to be a *-homomorphism.

Definition 4.45 The quadruple $(\mathcal{A}^+, \varphi^+, \mathcal{J}^+, i^+)$ is called *asymptotically complete* if for every $a \in \mathcal{A}$

$$\lim_{n\to\infty} ||(\mathcal{J}^+)^n i^+(a) - Q^+(\mathcal{J}^+)^n i^+(a)||_{\varphi^+} = 0 \ ,$$

where $Q^+ \colon \mathcal{A}^+ \to \mathcal{C}^+,\ Q^+(a \otimes c) := \varphi(a)\mathbb{1}_{\mathcal{C}^+}$.

Theorem 4.46 *Let $T \colon \mathcal{A} \to \mathcal{A}$ be a unital, completely positive operator with a synchronising convex decomposition and $\mathcal{J} \colon (\mathcal{A}, \theta_\infty) \to (\mathcal{A} \otimes \mathcal{C}, \theta_\infty \otimes \psi_\lambda),\ a \mapsto \sum_i T_i(a) \otimes \varepsilon_i$. Assume additionally that θ_∞ is faithful. Then $(\mathcal{A}^+, \varphi^+, \mathcal{J}^+, i^+)$ is asymptotically complete.*

Proof. As the operator T is synchronisable there exists a natural number n_0, such that $T^{n_0} = \lambda E_\theta + (1-\lambda)S$ for some state $\theta \in \mathcal{S}(\mathcal{A})$, some operator $S \in CP_u(\mathcal{A})$ and $\lambda \in (0, 1]$. Being interested into the asymptotic behaviour of T, we may consider T^{n_0} instead of T.
Assume the operator T possesses a convex decomposition of the form

$$T = \lambda E_\theta + (1-\lambda)S$$

with $\theta \in \mathcal{S}(\mathcal{A})$. Then $\mathcal{J}(a)$ is given by

$$\begin{aligned}\mathcal{J}(a) &= E_\theta(a) \otimes \varepsilon_1 + S(a) \otimes \varepsilon_2 \\ &= \theta(a)\mathbb{1} \otimes \varepsilon_1 + S(a) \otimes \varepsilon_2\end{aligned}$$

and by Lemma 4.2

$$\begin{aligned}
(\mathcal{J}^+)^n(i^+(a)) &= E_\theta(a) \otimes_{n-1} \mathbb{1}_\mathcal{C} \otimes \varepsilon_1 \otimes \mathbb{1}_{\mathcal{C}^+} \\
&+ E_{\theta\circ S}(a) \otimes_{n-2} \mathbb{1}_\mathcal{C} \otimes \varepsilon_1 \otimes \varepsilon_2 \otimes \mathbb{1}_{\mathcal{C}^+} \\
&\ \ \vdots \\
&+ E_{\theta\circ S^{n-1}}(a) \otimes \varepsilon_1 \otimes_{n-1} \varepsilon_2 \otimes \mathbb{1}_{\mathcal{C}^+} \\
&+ S^n(a) \otimes_n \varepsilon_2 \otimes \mathbb{1}_{\mathcal{C}^+}
\end{aligned}$$

and

$$\begin{aligned}
Q^+(\mathcal{J}^+)^n i^+(a) &= E_{\theta_\infty}(\mathcal{J}^+)^n i^+(a) \\
&= E_{\theta_\infty} E_\theta(a) \otimes_{n-1} \mathbb{1}_\mathcal{C} \otimes \varepsilon_1 \otimes \mathbb{1}_{\mathcal{C}^+} \\
&+ E_{\theta_\infty} E_{\theta\circ S}(a) \otimes_{n-2} \mathbb{1}_\mathcal{C} \otimes \varepsilon_1 \otimes \varepsilon_2 \otimes \mathbb{1}_{\mathcal{C}^+} \\
&\ \ \vdots \\
&+ E_{\theta_\infty} E_{\theta\circ S^{n-1}}(a) \otimes \varepsilon_1 \otimes_{n-1} \varepsilon_2 \otimes \mathbb{1}_{\mathcal{C}^+} \\
&+ E_{\theta_\infty} S^n(a) \otimes_n \varepsilon_2 \otimes \mathbb{1}_{\mathcal{C}^+} \\
&= (\mathcal{J}^+)^n(i^+(a)) \\
&+ E_{\theta_\infty} S^n(a) \otimes_n \varepsilon_2 \otimes \mathbb{1}_{\mathcal{C}^+} - S^n(a) \otimes_n \varepsilon_2 \otimes \mathbb{1}_{\mathcal{C}^+} \ .
\end{aligned}$$

Therefore

$$||(\mathcal{J}^+)^n(i^+(a)) - Q^+(\mathcal{J}^+)^n i^+(a)||_{\varphi^+} \leq 2(1-\lambda)^n |\theta_\infty(a)|^2$$

and from that the assertion follows. □

Unfortunately, this is only half of the reversion of Theorem 4.43, since the operator $\mathcal{J}$ is not a *-homomorphism. However, if we restrict ourselves to commutative W^*-algebras $\mathcal{A} = L^\infty(A, \Sigma_A, \mu)$ and $\mathcal{C} = L^\infty(C, \Sigma_C, \nu)$ for finite sets A and C, the operator $\mathcal{J}$ is a *-homomorphism if the operators of the convex decomposition are extremal in $CP_u(\mathcal{A}, \mathcal{A})$.

Lemma 4.47 *Let $T\colon L^\infty(A, \Sigma_A, \mu) \to L^\infty(A, \Sigma_A, \mu)$ be a unital, completely positive operator with convex decomposition $T(x) = \sum_{c\in C} \lambda_c T_c$ with $\lambda_c \in [0,1]$, $\sum_{c\in C} \lambda_c = 1$,* *such that $T_c \in CP_u(\mathcal{A}, \mathcal{A})$ is extremal for all $c \in C$. Then*

$$\mathcal{J}\colon (L^\infty(A,\mu), \varphi_\mu) \to (L^\infty(A,\mu) \otimes L^\infty(C,\lambda), \varphi_\mu \otimes \psi_\lambda),\ a \mapsto \sum_c T_c(a) \otimes \varepsilon_c$$

*is a *-homomorphism.*

Proof. Assume $A = \{1, \dots, n\}$. Let $T_c \in CP_u(\mathcal{A}, \mathcal{A})$ be extremal. Then $T_c(f) = (f_{\pi_c(1)}, \dots, f_{\pi_c(n)})$ for all $f = (f_1, \dots, f_n) \in L^\infty(A, \Sigma_A, \mu)$ and a mapping $\pi_c\colon A \to A$. Consequently, we can verify multiplicativity for elements $g, f \in L^\infty(A, \Sigma_A, \mu)$ easily

$$\begin{aligned} T_c(g \cdot f) =& ((g \cdot f)_{\pi_c(1)}, \dots, (g \cdot f)_{\pi_c(n)}) \\ =& (g_{\pi_c(1)} \cdot f_{\pi_c(1)}, \dots, g_{\pi_c(n)} \cdot f_{\pi_c(n)}) = T_c(g) \cdot T_c(f)\ . \end{aligned}$$

□

Corollary 4.48 *Let $T\colon L^\infty(A, \Sigma_A, \mu) \to L^\infty(A, \Sigma_A, \mu)$ be an operator with synchronising convex decomposition $T(x) = \sum_{i\in I} \lambda_i T_i(x)$, such that the operators $T_i \in CP_u(L^\infty(A), L^\infty(A))$ are extremal for all $i \in I$. Then $\mathcal{J}\colon (L^\infty(A,\mu), \varphi_\mu) \to (L^\infty(A,\mu) \otimes L^\infty(C,\nu), \varphi_\mu \otimes \psi_\lambda)$ is an asymptotically complete random variable.*

The equivalence a) ⇔ b) in the following theorem was stated and proved in [GKL06]. We add a further equivalence.

Theorem 4.49 *Let A, C be finite sets with measures μ and λ, respectively, and $\gamma\colon A \times C \to A$ a measure preserving transformation. Further, let*

$$J\colon (L^\infty(A,\mu), \varphi_\mu) \to (L^\infty(A,\mu) \otimes L^\infty(C,\lambda), \varphi_\mu \otimes \psi_\nu)$$

be the random variable associated to γ. Then the following assertions are equivalent:

a) The graph G_γ possesses a synchronising word.

b) The random variable J is asymptotically complete.

c) The convex decomposition $P_{\psi_\lambda} J = \sum_{c\in C} \lambda(c) T_c$ with extremal transition matrices $T_c = (\delta_{\gamma(i,c),j})_{i,j}$, $c \in C$, is synchronising.

In particular, the random variable can be regained directly from the decomposition of T_{ψ_λ} given in c) via $J(x) = \sum_{c\in C} T_c(x) \otimes \varepsilon_c$.

Proof. **a) ⇒ c)** Let the graph G_γ posses the synchronising word $c_1 \cdots c_n$. For each colour $c \in C$ we put $T_c = (\delta_{\gamma(i,c),j})_{i,j}$ to be the transition matrix that is associated to c canonically. In the set of all stochastic matrices the matrices T_c, $c \in C$, are extremal. We define another transition matrix

$$T_\gamma := \sum_{c\in C} \lambda(c) T_c \ ,$$

such that the matrices T_c provide an extremal convex decomposition with convex coefficients $\lambda(c)$. Note that the i,j-th entry $(T_\gamma)_{i,j}$ is equal to the probability $\lambda(\{c \in C \colon \gamma(i,c) = j\})$. Therefore, $T_\gamma = P_\psi J$. Under the assumption that the target vertex of the synchronising word $c_1 \cdots c_n$ is $a_0 \in A$, direct computation yields

$$T_{c_1} \cdots T_{c_n} = (\delta_{\gamma(\gamma(\cdots(\gamma(i,c_n),c_{n-1}),\cdots),c_1),j})_{i,j} = (\delta_{a_0,j})_{i,j} \ .$$

Hence, $\sum_{c\in C} \lambda(c)T_c$ is a synchronising convex decomposition.

c) ⇒ a) As $T = \sum_{c\in C} \lambda(c)T_c$ is a synchronising convex decomposition, there exists a product $T_{c_1} \cdots T_{c_n}$ such that $T_{c_1} \cdots T_{c_n}$ is a projection on a single edge, say $a_0 \in A$. It is understood, that this translates into $c_1 \cdots c_n$ being a synchronising word for the graph G_γ with target edge $a_0 \in A$.

b) ⇒ c) For the proof of this implication we refer to [GKL06].

c) ⇒ b) The mapping

$$\mathcal{J} \colon (L^\infty(A,\mu), \varphi_\mu) \to (L^\infty(A,\mu) \otimes L^\infty(C,\lambda,\varphi_\mu \otimes \psi_\lambda),\ a \mapsto \sum_i T_i(a) \otimes \varepsilon_i$$

is an asymptotically complete random variable by Corollary 4.48. □

4.5 Perspectives

Already for c-graphs with countably many vertices the existence of a finite synchronising word is a very restrictive condition, and usually we will not be able to find such a graph, neither by recolouring nor by splitting edges. As an appropriate substitute F. Haag [Haa06] proposed a definition for a generalisation of the concept of a synchronising word for graphs with a countable set of vertices.

Definition 4.50 Let $G = (V, E)$ be a c-graph with a countable set of vertices V and set of colours C. The graph G is said to be *synchronising* if for every finite $V_0 \subset V$ there exists a finite sequence of $c_1, \dots, c_n$ of colours $c_i \in C$, $i = 1, \dots, n$, such that for every sequence $e_1, \dots, e_n$ of edges $e_i \in E$, $i = 1, \dots, n$, with $t(e_i) = s(e_{i+1})$ for all $i = 1, \dots, n-1$ and $c(e_i) = c_i$ for all $i = 1, \dots, n$ the target vertex $t(e_n)$ is the same.

Justifying his definition in hindsight, F. Haag [Haa06] has proven an analogous theorem to Theorem 2.10.

Theorem 4.51 *Let $\gamma\colon (A \times C, \mu \otimes \nu) \to (A, \mu)$ be a measure preserving mapping. The following assertions are equivalent:*

a) The graph G_γ is synchronising.

b) The random variable J is asymptotically complete.

In the line of this equivalence we would like to give a generalised definition of a synchronisable operator, such that it adds as a further equivalence the existence of a synchronisable transition operator T_ψ.
Presumably closely related to the above question is the idea of a generalisation that may bear a name like ‘aproximately synchronisable’ operator. What we have in mind is a synchronising convex decomposition with synchronising words whose occurence guarantees that the entire state space is contracted at most to some ball of predetermined radius. Instead of a random family θ_i of states $\theta_i \in \mathcal{S}(\mathcal{A})$ we obtain a random family B_i of balls $B_i \subset \mathcal{S}(\mathcal{A})$. The convex combination of those random balls with respect to the empirical measure encloses the invariant state.

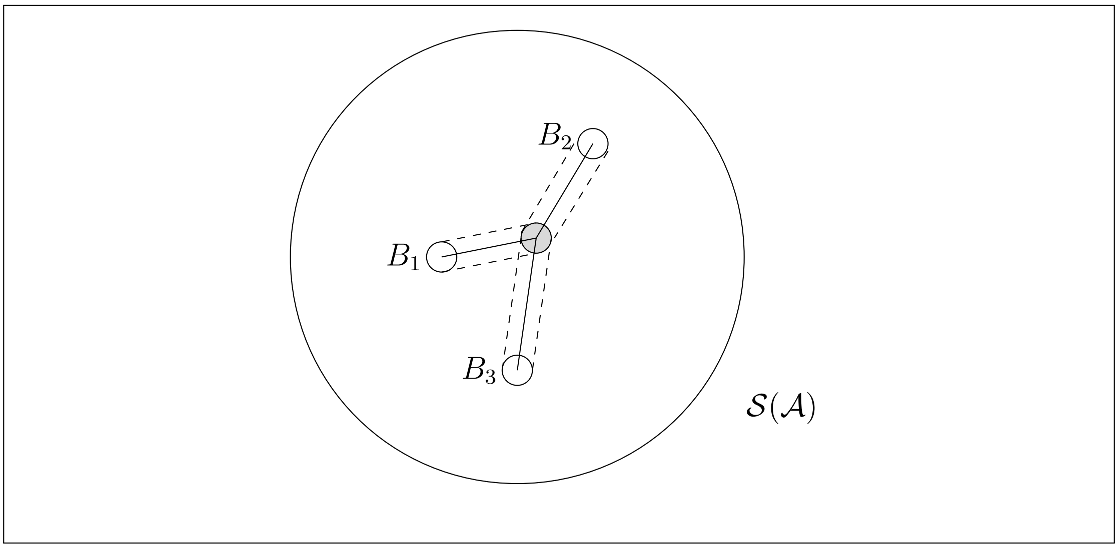


Figure 4.2: Convex combination of balls of states

CHAPTER 5

EXAMPLES

In this section we would like to point out how synchronising convex decompositions of operators may be used to approximate the invariant state algorithmically or to gain samples of states that are distributed like the invariant state. This may serve as another justification to consider our approach as a non-commutative version of J. D. Propp's and D. B. Wilson's Coupling from the Past algorithm.

5.1 Numerical Approach

Assume $\Omega = \{0,1\}$ with probability measure $\nu = (\lambda_0, \lambda_1)$. We define $\Gamma = \{0, 01, 011, 0111, \ldots\}$ to be the set of all finite words that begin with the element 0 and, apart from that, are made up of the element 1. In slight abuse of notation we denote a word $0\underbrace{1\cdots1}_{k\times}$ in Γ by 01^k. We assign to each word the measure $\hat{\nu}(01^k) = \lambda_0\lambda_1^k$. For later convenience we sort Γ by increasing word length or decreasing measure.

The following assertions are true for a general finite set $\Omega := \{0, \ldots, n\}$ and

$$\Gamma := \big\{ i_0 i_1 \cdots i_k \colon i_0 \in \{0, \ldots, m\}, \\ i_j \in \{n-m, \ldots, n\}, j = 1, \ldots, k, \, k \in \mathbb{N}_0 \big\} .$$

For convenience, however, we only prove them for the special case $\Omega = \{0, 1\}$; the general proofs are analogous and provide no further insight. Apart from Proposition 5.1 and 5.2 the set Γ may be considered as set of finite words of the above kind without further changes to the assertions.
Assume $X \colon (\Omega, \Sigma, \mu) \to \Gamma$ to be a random variable, such that the image measure $\mu X^{-1} = \hat{\nu}$. From X we construct a further random variable $\delta_X \colon (\Omega, \Sigma, \mu) \to \mathcal{M}(\Gamma)$ with values in the probability measures $\mathcal{M}(\Gamma)$ on Γ by $\delta_X(\omega) := \delta_{X(\omega)}$ where $\delta_x \in \mathcal{M}(\Gamma)$ is the point measure for every $x \in \Gamma$. We consider $\mathcal{M}(\Gamma)$ as a Banach space with the norm induced by the variation norm defined on the space of all measures on Γ:

$$||m||_{var} := \sup_{Z \in \mathcal{Z}} \sum_{A \in Z} m(A) \quad \text{for each } m \in \mathcal{M}(\Gamma) ,$$

where $\mathcal{Z}$ is the set of all partitions of Γ.

Proposition 5.1 *If $\delta_X \colon (\Omega, \Sigma, \mu) \to \mathcal{M}(\Gamma)$ denotes the random variable we have defined above, then*

$$E(\delta_X) = \sum_{k=1}^{\infty} \lambda_0 \lambda_1^k \delta_{1^k 0} = \hat{\nu} \ .$$

(For a general finite set Γ: $E(\delta_X) = \hat{\nu}$.)

Proof. The proof is immediate. □

From now on we are going to refer to $E(\delta_X)$ as $\hat{\nu}$ and consider the centralised random variable $\overline{\delta}_X = \delta_X - \hat{\nu}$. Let $(X_i)_{i \in \mathbb{N}_0}$, $X_0 = X$, be a sequence of independent, identically distributed (i.i.d.) random variables. The corresponding sequence of random variables $(\delta_{X_i})_{i \in \mathbb{N}_0}$ is i.i.d. as well.

Proposition 5.2 *If* $\overline{\delta}_X\colon (\Omega, \Sigma, \mu) \to \mathcal{M}(\Gamma)$ *denotes the random variable we have defined above and* $||\overline{\delta}_X||_{var}$ *the random variable* $||\overline{\delta}_X||_{var}\colon (\Omega, \Sigma, \mu) \to [0,1],\ \omega \mapsto ||\overline{\delta}_{X(\omega)}||_{var}$. *Then the expectation value of* $||\overline{\delta}_X||_{var}$ *is*

$$E(||\overline{\delta}_X||_{var}) = 2\frac{\lambda_1}{1+\lambda_1} < \infty \ .$$

(For a general finite set Γ*:* $E(||\overline{\delta}||_{var}) < \infty$*.)*

Proof. It is obvious that we can identify the elements in Γ with natural numbers. Our summation index simplifies considerably, and we obtain:

$$\begin{aligned} E(||\overline{\delta}_X||_{var}) = \sum_{i=0}^{\infty} ||\overline{\delta}_i||_{var}\hat{\nu}_i &= \sum_{i=0}^{\infty} \hat{\nu}_i \frac{1}{2}(|1-\hat{\nu}_i| + \sum_{j\neq i} \hat{\nu}_j) = \sum_{i=0}^{\infty} \hat{\nu}_i(1-\hat{\nu}_i) \\ &= \sum_{i=0}^{\infty} \hat{\nu}_i - \sum_{i=0}^{\infty} \hat{\nu}_i^2 = 1 - \lambda_0^2 \sum_{i=0}^{\infty} (\lambda_1^2)^i = 2\frac{\lambda_1}{1+\lambda_1} \ . \end{aligned}$$

□

Hence, the sequence $(\overline{\delta}_{X_i})_{i\in\mathbb{N}_0}$ meets all requirements for the strong law of large numbers on Banach spaces.

Theorem 5.3 *[LT91] Let* $(Y_n)_{n\in\mathbb{N}_0}$ *be a sequence of i.i.d. Borel random variables distributed as* Y_0 *with values in a separable Banach space* E*. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} Y_i = 0 \quad \textit{almost surely}$$

if and only if $E(||Y_0||) < \infty$ *and* $E(Y_0) = 0$*.*

Corollary 5.4 *If* $(\delta_{X_k})_{k\in\mathbb{N}_0}$ *denotes the sequence of the random variables we defined above, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \delta_{X_k} = \hat{\nu} \quad \textit{almost surely} \ .$$

If $T\colon \mathcal{A} \to \mathcal{A}$ is a synchronisable operator with decomposition $T = \lambda_0 E_\varphi + \lambda_1 S$, the words in the set Γ represent the products $E_\varphi S^k$, i.e. the word 01^k corresponds to the operator $E_\varphi S^k$. In that sense we may associate every measure in $\nu \in \mathcal{M}(\Gamma)$ with an operator $\mathcal{E}(\nu)$ on $\mathcal{A}$ by

$$\mathcal{E}\colon \mathcal{M}(\Gamma) \to \mathcal{C}P_u(\mathcal{A}),\ \nu \mapsto \sum_{k=1}^{\infty} \nu_k E_\varphi S^k \ ,$$

or for an arbitrary state $\theta \in \mathcal{S}(\mathcal{A})$, with a state on $\mathcal{A}$ by

$$\mathcal{M}(\Gamma) \to \mathcal{S}(\mathcal{A}),\ \nu \mapsto \theta \circ \mathcal{E}(\nu) \ .$$

In particular, the measure $\hat{\nu}$ is mapped on the conditional expectation E_{θ_∞} on the invariant state θ_∞ of the operator:

Proposition 5.5 *If $T\colon \mathcal{A} \to \mathcal{A}$ is a unital, completely positive and synchronisable operator on a C^*-algebra with identity that possesses a synchronising convex decomposition of the form $T = \lambda_0 E_\varphi + \lambda_1 S$ for a state $\varphi \in \mathcal{S}(\mathcal{A})$, an operator $S \in CP_u(\mathcal{A}, \mathcal{A})$ and $\lambda_0, \lambda_1 \in (0,1)$ with $\lambda_0 + \lambda_1 = 1$, then*

$$\mathcal{E}(\hat{\nu}) = \sum_{k=0}^{\infty} \lambda_0 \lambda_1^k E_\varphi S^k = E_{\theta_\infty} \ .$$

Proof. First we show that the operator $\mathcal{E}(\frac{1}{n}\sum_{i=0}^{n-1} \delta_{X_i})$ converges to E_{θ_∞} uniformly:

$$\begin{aligned}
&||\mathcal{E}(\frac{1}{n}\sum_{i=0}^{n-1} \delta_{X_i}) - E_{\theta_\infty}|| = ||\sum_{k=0}^{\infty} \frac{1}{n} \sum_{i=0}^{n-1} \delta_{X_i}(k) E_\varphi S^k - E_{\theta_\infty}|| \\
&= ||\frac{1}{n}\sum_{i=0}^{n-1} \lambda_0 \lambda_1^k E_\varphi S^k - E_{\theta_\infty}|| = ||\sum_{i=0}^{n-1} \lambda_0 \lambda_1^k E_\varphi S^k - E_{\theta_\infty}|| \\
&\leq ||\sum_{i=0}^{n-1} \lambda_0 \lambda_1^k E_\varphi S^k - \sum_{k=0}^{\infty} \lambda_0 \lambda_1^k E_\varphi S^k|| = ||\sum_{k=n}^{\infty} \lambda_0 \lambda_1^k E_\varphi S^k|| \\
&\leq ||\lambda_0 \lambda_1^n \sum_{k=n}^{\infty} \lambda_1^k|| = \lambda_1^n \ .
\end{aligned}$$

As $\mathcal{E}\colon \mathcal{M}(\Gamma) \to \mathcal{CP}_u(\mathcal{A})$ is linear and bounded

$$||\mathcal{E}(\hat{\nu}) - E_{\theta_\infty}|| \leq ||\mathcal{E}(\hat{\nu}) - \mathcal{E}(\frac{1}{n}\sum_{i=0}^{n-1}\delta_{X_i})|| + ||\mathcal{E}(\frac{1}{n}\sum_{i=0}^{n-1}\delta_{X_i}) - E_{\theta_\infty}||$$
$$= ||\hat{\nu} - \frac{1}{n}\sum_{i=0}^{n-1}\delta_{X_i}|| + \lambda_1^n \ .$$

By Corollary 5.4 the last term converges to 0 almost surely. □

Therefore, we can produce random samples of states $\varphi_i \in \mathcal{S}(\mathcal{A})$ that are distributed like θ_∞ with respect to the measure.

5.2 First Example

Let $\mathcal{A}$ be the matrix algebra M_2 and $T\colon M_2 \to M_2$ a completely positive normal operator given by $T(x) = \frac{1}{4}E_{\varphi_+}(x) + \frac{1}{4}E_{\varphi_-}(x) + \frac{1}{2}S(x)$ where

$$S(x) = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} x \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

and $\varphi_+, \varphi_- \in \mathcal{S}(M_2)$ are states on M_2 having density matrices $\rho_+ = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $\rho_- = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, respectively. Explicitly, the operator $T\colon M_2 \to M_2$ is given by

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \mapsto \frac{1}{4}\begin{pmatrix} 5x_{11} - 2x_{12} - 2x_{21} + 5x_{22} & 5x_{11} + 2x_{12} + 2x_{21} - x_{22} \\ 5x_{11} - 2x_{12} + 2x_{21} - x_{22} & 5x_{11} + 2x_{12} + 2x_{21} + 5x_{22} \end{pmatrix} \ .$$

Then T is a synchronisable operator with the trace state as invariant state θ_∞. We interpret the convex coefficients of the decomposition of T as a measure $\nu = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$ on $\Omega := \{0, 1, 2\}$. From it we construct another measure $\hat{\nu}$ on $\Gamma := \{2^k 0\colon k \in \mathbb{N}_0\} \cup \{2^k 1\colon k \in \mathbb{N}_0\}$ by $\hat{\nu}(2^k 1) = \frac{1}{4}(\frac{1}{2})^k$ and $\hat{\nu}(2^k 0) = \frac{1}{4}(\frac{1}{2})^k$. The mapping $\mathcal{E}\colon \Gamma \to \mathcal{CP}_u(M_2)$ relates those words to products of operators of the convex decomposition. By an easy calculation we see that among the products

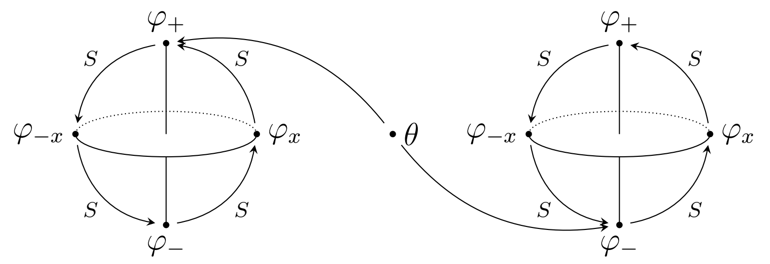


Figure 5.1: Process on the states $\theta, \varphi_+, \varphi_-, \varphi_x$ and φ_{-x}

occur four different operators as results only:

$$\begin{aligned} E_{\varphi_+}S^{4k} &= E_{\varphi_+} & E_{\varphi_-}S^{4k} &= E_{\varphi_-} \\ E_{\varphi_+}S^{4k+1} &= E_{\varphi_x} & E_{\varphi_-}S^{4k+1} &= E_{\varphi_{-x}} \\ E_{\varphi_+}S^{4k+2} &= E_{\varphi_-} & E_{\varphi_-}S^{4k+2} &= E_{\varphi_+} \\ E_{\varphi_+}S^{4k+3} &= E_{\varphi_{-x}} & E_{\varphi_-}S^{4k+3} &= E_{\varphi_x} \ , \end{aligned}$$

where $\varphi_x, \varphi_{-x} \in \mathcal{S}(M_2)$ are the states with density matrices $\rho_x = \frac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ and $\rho_{-x} = \frac{1}{2}\left(\begin{smallmatrix} 1 & -1 \\ 1 & -1 \end{smallmatrix}\right)$. Instead of the four conditional expectations we can consider the four states $\varphi_-, \varphi_+, \varphi_x$ and φ_{-x}. Thought as outputs of a random event they are distributed according to the measure $\mu \colon (\mathcal{S}(M_2), \mathcal{B}(\mathcal{S}(M_2))) \to [0,1]$ with

$$\mu(\cdot) := \frac{1}{3}(\delta_{\varphi_+} + \delta_{\varphi_-}) + \frac{1}{6}(\delta_{\varphi_x} + \delta_{\varphi_{-x}}) \ ,$$

where δ_φ is the point measure of the state $\varphi \in \mathcal{S}(M_2)$. Hence, this is an example of a transition operator which behaves very classical, in that sense that by coupling it to a process on Ω we are dealing with a process on finitely many states only.

5.2.1 Simulation

We produce 100000 random words $w_i \in \Gamma$ and the corresponding samples of states $\{ \theta_i \}_i$, which are distributed according to μ. From the family of samples $\{ \theta_i \}_i$ we

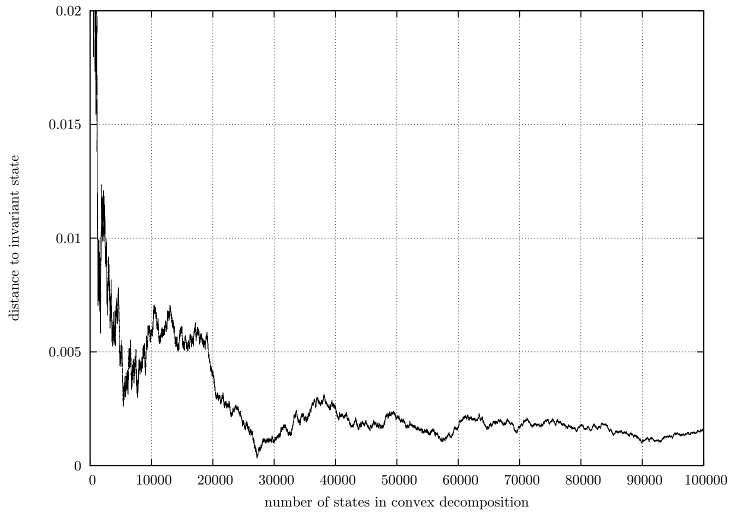


Figure 5.2: Single trajectory in $\mathcal{S}(M_2)$

build the arithmetic means

$$\Theta_n := \frac{1}{n} \sum_{k=0}^{n-1} \theta_k \ .$$

The means Θ_n converge to θ_∞ almost surely. To illustrate that we calculate the euclidean distance $||\Theta_n - \theta_\infty||$ to the invariant state θ_∞ (Figure 5.2).
In another experiment we try to get an idea of the expected euclidean distance. As before we produce samples of states $\{ \theta_i \}_i$, $i = 1, \ldots, 1000$, and their means Θ_i. This time, however, we repeat this 1000 times and obtain 1000 families of samples $\left\{ \{ \theta_i^j \}_i : j = 1, \ldots, 1000 \right\}$, means $\left\{ \{ \Theta_i^k \}_i : k = 1, \ldots, 1000 \right\}$ and euclidean distances to the invariant state θ_∞. By taking the arithmetic mean of the distances we smooth individual fluctuations (Figure 5.3).

5.3 Truncated Micromaser

The following example is gained by a simplification of the stochastic model used to describe the micromaser, introduced in section 1.2.13. We truncate it, such

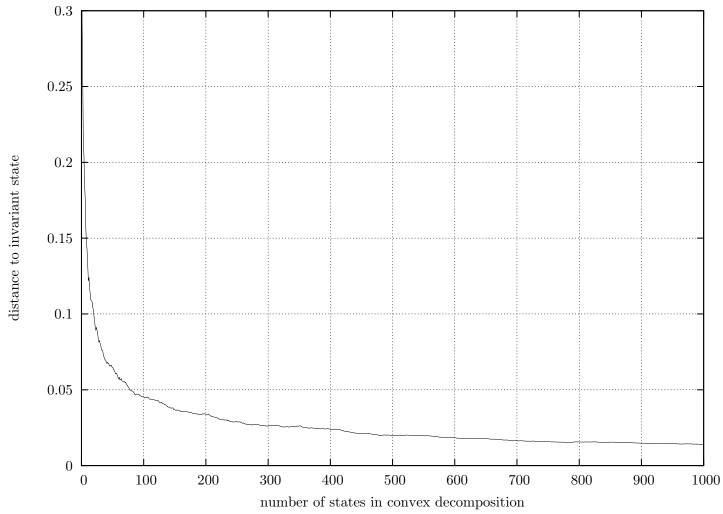


Figure 5.3: Expected distance in $\mathcal{S}(M_2)$

that the Hilbert spaces are both finite dimensional. In the description of the real micromaser the Hilbert space representing the standing wave in the cavity is infinite dimensional. However, in that case neither the transition operator nor any power can be decomposed into a synchronising convex decomposition. Let $\mathcal{A} = M_n$, $\mathcal{C} = M_2$ and $J\colon \mathcal{A} \to \mathcal{A} \otimes \mathcal{C}$ be the random variable defined by $J\colon M_n \to M_n \otimes M_2,\ x \mapsto u^*(x \otimes \mathbb{1})u$ with a unitary operator $u \in M_n \otimes M_2$. Originating from the mathematical model of the micromaser the unitary operator u is given by

$$u := \begin{pmatrix} a_+ & s^*b \\ bs & a \end{pmatrix} \in M_2(M_n)$$

where a, a_+, b and s are $n \times n$-matrices

$$a := \begin{pmatrix} 1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix}, \qquad a_+ := \begin{pmatrix} \alpha_n & & & \\ & \ddots & & \\ & & \alpha_2 & \\ & & & 1 \end{pmatrix}$$

$$b := \begin{pmatrix} 0 & & & \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_n \end{pmatrix}, \qquad s := \begin{pmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}$$

with $0 \leq \alpha_k \leq 1$, $\alpha_1 = 1$ and $\beta_k = i\sqrt{1-\alpha_k^2}$. The random variable J may be written as

$$\begin{aligned} J(x) &= \begin{pmatrix} a_+ & s^*b^* \\ b^*s & a \end{pmatrix} \cdot \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} \cdot \begin{pmatrix} a_+ & s^*b \\ bs & a \end{pmatrix} \\ &= \begin{pmatrix} T_1(x) & T_2(x) \\ T_3(x) & T_4(x) \end{pmatrix} \end{aligned}$$

with

$$\begin{aligned} T_1(x) =& a_+xa_+ + s^*b^*xbs \\ T_2(x) =& a_+xs^*b + s^*b^*xa \\ T_3(x) =& b^*sxa_+ + axbs \\ T_4(x) =& b^*sxs^*b + axa \ . \end{aligned}$$

In [GKL06] R. Gohm, B. Kümmerer and T. Lang have proven a criterion that assures that J is asymptotically complete.

Theorem 5.6 *With $\mathcal{A}, \mathcal{C}$ and J as above the following conditions are equivalent:*

a) $\alpha_i \neq 1$ for $2 \leq i \leq n$,

b) J is asymptotically complete.

Corollary 5.7 *If $\alpha_i \neq 1$ for $2 \leq i \leq n$, then the transition operators $T_\psi := P_\psi J$ posses a synchronising convex decomposition for all normal states $\psi \in \mathcal{S}(\mathcal{A})$.*

Proof. The assertion follows from Theorem 4.43. □

For convenience we choose a pleasant set of parameters a_k, although we could pick any set fulfilling $\alpha_k \neq 1$ for $2 \leq k \leq n$. Our choice is $a_1 = 1$ and $a_k = 0$ for all $k > 1$. If $\psi \in \mathcal{S}(M_2)$ is a normal state on M_2 with density

$$\rho_\psi = \begin{pmatrix} \lambda & i\sqrt{\lambda(1-\lambda)}\,\zeta \\ -i\sqrt{\lambda(1-\lambda)}\,\overline{\zeta} & 1-\lambda \end{pmatrix}$$

where $\lambda \in [0,1]$ and $\zeta \in \mathbb{C}, |\zeta| = 1$, the transition operator T_ψ can be presented by determining its effect on matrix units (Figure 5.4). The unusual parametrisation is due to the advantage of having neat expressions as 'transition' coefficients.
In the following we perform simulations for the special cases $\mathcal{A} = M_3$ and $\mathcal{A} = M_{10}$.

5.4 Second Example

Let $\mathcal{A}$ be the matrix algebra M_3 and $\psi \in \mathcal{S}(M_2)$ be the normal state with density matrix $\rho_\psi = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ with $\lambda_1, \lambda_2 \in [0,1]$ $\lambda_1 + \lambda_2 = 1$. Let further $u = \begin{pmatrix} a_+ & s^*b \\ bs & a \end{pmatrix}$ be the unitary operator given by the set of matrices

$$a_+ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad a = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad b = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad s = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \ .$$

The invariant state θ_∞ of the transition operator T_ψ has the density matrix $\rho = \frac{4}{19}\begin{pmatrix} \frac{9}{4} & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Direct computation shows that the third power of the transition

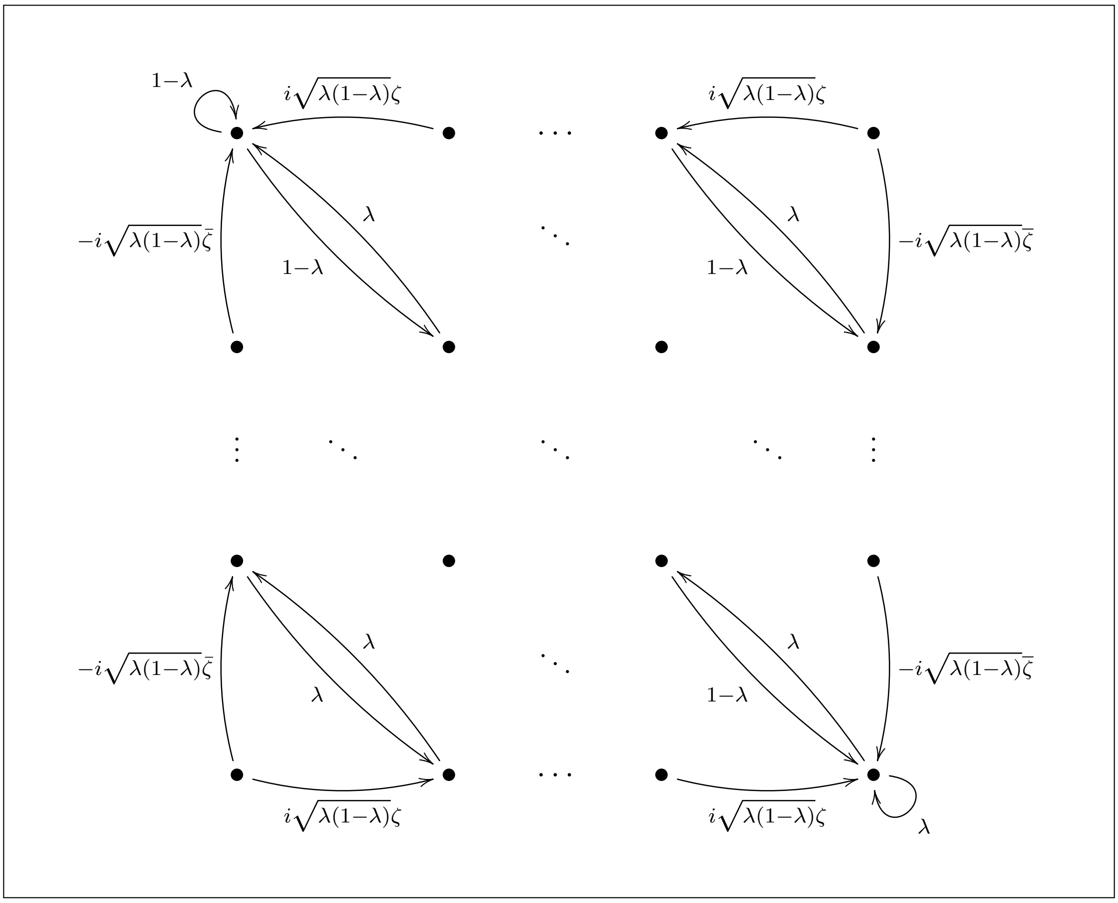


Figure 5.4: Transitions of T_ψ

operator $T_\psi := P_\psi J$ is full, indeed

$$\begin{aligned}(T_\psi)^3(x) =& (\lambda_1 3 + \lambda_1\lambda_2^2)E_{11}xE_{11} + (\lambda_2^3 + \lambda_1^2\lambda_2)E_{21}xE_{12} + \lambda_1\lambda_2^2 E_{31}xE_{13} \\ &+ (\lambda_1^2\lambda_2 + \lambda_1\lambda_2^2)(E_{12}xE_{21} + E_{22}xE_{22} + E_{32}xE_{23}) \\ &+ \lambda_1^2\lambda_2 E_{13}xE_{31} + (\lambda_1^3 + \lambda_1^2\lambda_2)E_{23}xE_{32} + (\lambda_2 3 + \lambda_1^2\lambda_2)E_{33}xE_{33} \\ &+ \lambda_1^2\lambda_2(E_{12} + E_{23})x(E_{21} + E_{32}) + \lambda_1\lambda_2^2(E_{21} + E_{32})x(E_{12} + E_{23}) \ .\end{aligned}$$

For the three states $\varphi_1, \varphi_2, \varphi_3 \in \mathcal{S}(\mathcal{A})$ with density matrices

$$\rho_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ , \quad \rho_2 = \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ , \quad \rho_3 = \frac{4}{19}\begin{pmatrix} \frac{10}{4} & 0 & 0 \\ 0 & \frac{5}{4} & 0 \\ 0 & 0 & 1 \end{pmatrix} \ ,$$

respectively, we determine a positive real number $\alpha_i \in \mathbb{R}^+$, $i = 1, 2, 3$, such that $\alpha_i E_{\varphi_i} < T_\psi$ for each $i = 1, 2, 3$. We have described that procedure in Section 4.2.1. The terms on the right-hand side yield a concrete representation of the operator T^3. Already the terms $\{\, E_{ij} \colon i, j = 1, 2, 3 \,\}$ in the first three lines span M_3. Hence, we ignore the last two in the fourth line and save us some complication. The knowledge of the coefficients of the former terms suffices to determine a minor of the actual Radon-Nikodym density $\mathbb{1} \otimes \Lambda$. We have chosen the three factors $\alpha_1 = \frac{18}{125}, \alpha_2 = \frac{38}{125}$ and $\alpha_3 = \frac{114}{625}$. Now T_ψ^3 can be decomposed into

$$T_\psi^3 = \alpha_i E_{\varphi_i} + (1 - \alpha_i) R_i$$

for each state $\varphi_i,\ i = 1, 2, 3$, where the remainder $R_i = \frac{1}{1-\alpha_i}(T_\psi^3 - \alpha_i E_{\varphi_i})$ is a unital and completely positive operator, too.

5.4.1 Simulation

As before we produce 100000 random words $w_i \in \Gamma$, the corresponding samples of states and their arithmetic means $\{\, \Theta_i \,\}_i$. We calculate the euclidean distance to the invariant state θ_∞ and illustrate the convergence in Figure 5.5.
Then, for each decomposition $T_\psi^3 = \alpha_i E_\varphi + (1 - \alpha_i) R_i$, $i = 1, 2, 3$, we produce 1000

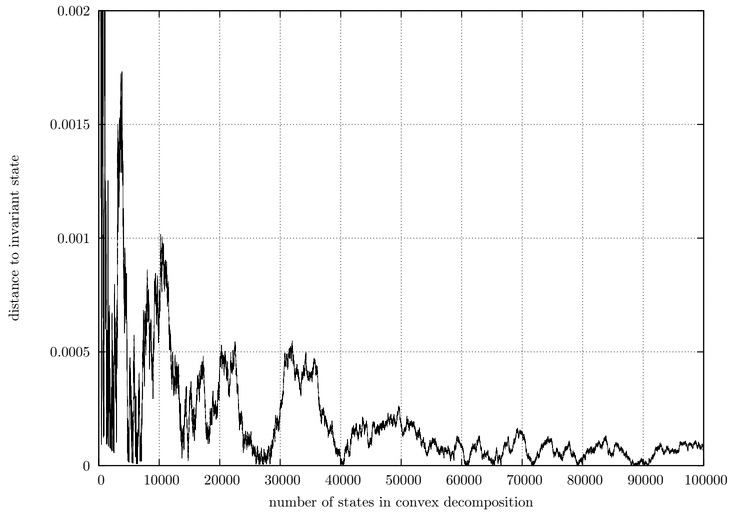


Figure 5.5: Single trajectory in $\mathcal{S}(M_3)$

families of states, their means and euclidean distances. In Figure 5.6 we present the mean of the euclidean distances for all three decompositions. Additionally we examined the convergence of $n^\kappa(\frac{1}{n}\sum_{i=1}\theta_i - \theta_\infty)$ for factors $\kappa \in (0,1]$. Some results for $\kappa_1 = \frac{2}{5}, \kappa_2 = \frac{1}{2}$ and $\kappa_3 = \frac{3}{5}$ are presented in Figure 5.7.

5.5 Third Example

This time we choose $\mathcal{A} = M_{10}$ and $\psi \in \mathcal{S}(M_2)$ with density matrix $\rho_\psi = \left(\begin{smallmatrix} \lambda & 0 \\ 0 & 1-\lambda \end{smallmatrix}\right)$. Let $\alpha_1 = 1$ and $\alpha_k = 0$ for $1 < k \leq 10$. The four matrices a_+, a, b and s are chosen just as in the second example except of their larger dimensions. Analogously to the transition operator on M_3 before T_ψ^{10} is full. However, this time the density Λ has to be a 10×10-matrix, such that direct computation is rather laborious. But without a lot of computation we see that the terms $E_{10,1}$ and $E_{1,10}$ only occur as product $E_{10,1} = a_+(bs)^9$ and $E_{1,10} = a(s^*b)^9$ with smallest coefficients $\lambda(1-\lambda)^9$ and $(1-\lambda)\lambda^9$, respectively. Either $\lambda(1-\lambda)^9$ or $(1-\lambda)\lambda^9$ is the smallest coefficient. Therefore, if $\varphi \in \mathcal{S}(M_{10})$ is the trace state, some calculation shows $\alpha E_\varphi \leq T_\psi$ for

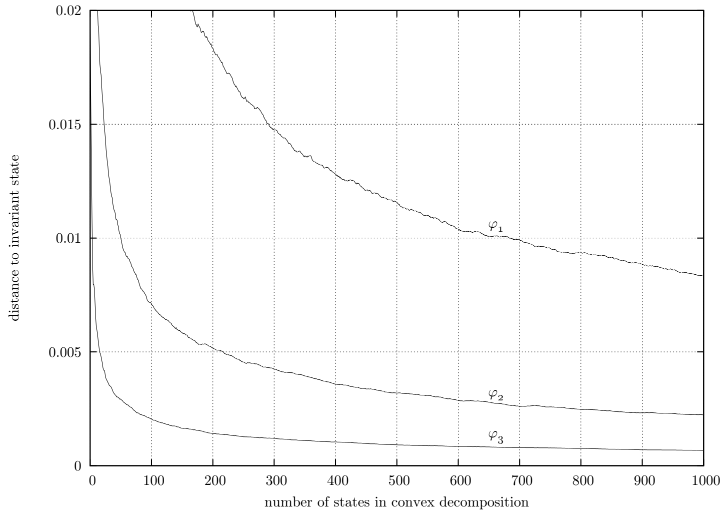


Figure 5.6: Expected distance in $\mathcal{S}(M_3)$

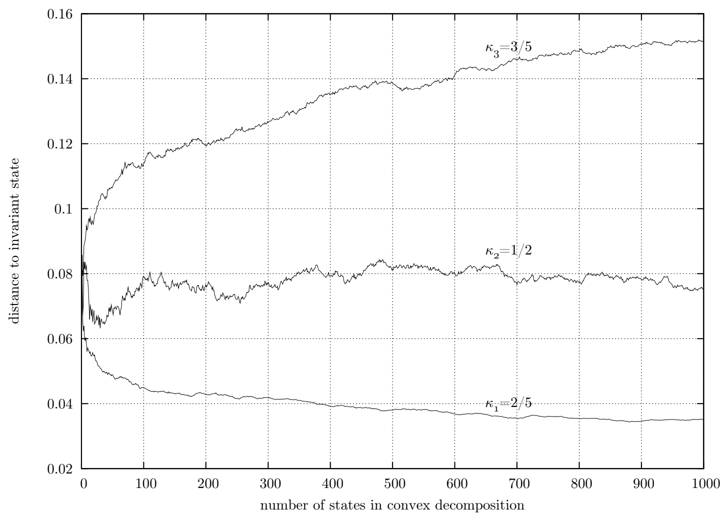


Figure 5.7: Single trajectories for different factors $\kappa_1 = \frac{2}{5}, \kappa_2 = \frac{1}{2}$ and $\kappa_3 = \frac{3}{5}$

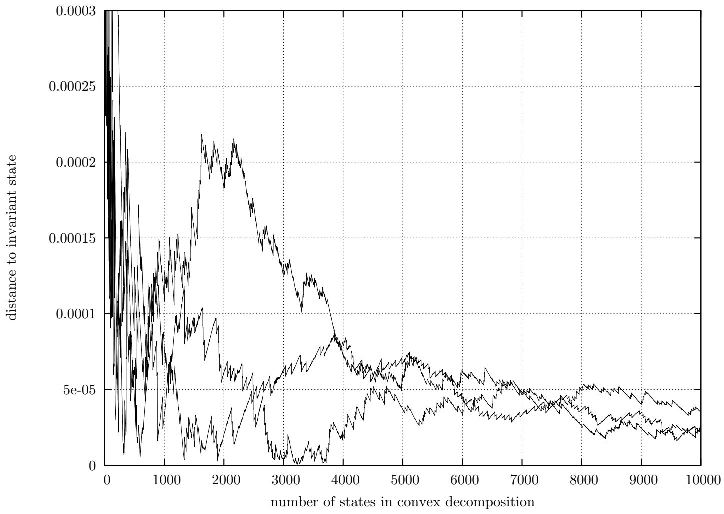


Figure 5.8: Three trajectories in $\mathcal{S}(M_{10})$ with identical parameters

$\alpha \leq 100 \min(\lambda(1-\lambda)^9, (1-\lambda)\lambda^9)$.

5.5.1 Simulation

We produce 10000 random samples of states $\{\, \theta_i \,\}_i$ and calculate their arithmetic means $\Theta_i = \frac{1}{n} \sum_{k=0}^{n-1} \theta_k$. Instead of the euclidean distance between the arithmetic means Θ_i and the invariant state θ_∞, we determine the distance between Θ_i and $\Theta_i \circ T$. In Figure 5.8 we present the distances of three numerical experiments with identical parameters.

5.6 Résumé

As we see from the figures clearly the arithmetic means become more and more stabilised with increasing number of states. The reason for this is that the arithmetic mean of n states $\theta_i,\ i = 1, \dots, n$ may be regarded as the convex combination of the arithmetic means of the first $n-1$ states and the n-th state

$$\frac{1}{n}\sum_{i=1}^{n}\theta_i = \frac{n-1}{n}\left(\frac{1}{n-1}\sum_{i=1}^{n-1}\theta_i\right) + \frac{1}{n}\theta_n \ .$$

Not only for practical reasons would it be interesting to know the rate of the convergence to the invariant state, that means to have an upper bound for expressions of the form

$$P\Big(||\frac{1}{n}\sum_{i=0}^{n-1}\theta_i - \theta_\infty|| < \varepsilon\Big) \quad \text{or} \quad \mathbb{E}\Big(||\frac{1}{n}\sum_{i=0}^{n-1}\theta_i - \theta_\infty|| < \varepsilon\Big) \ .$$

Using the Central Limit Theorem the first probability may be estimated at least pointwisely. Recall Section 5.1. Consider the random variable $(\Omega, \Sigma, \mu) \to CP_u(\mathcal{A}, \mathcal{A}),\ \omega \mapsto \mathcal{E}(\delta_X)$. Then for every $a \in \mathcal{A}_+$, $||a|| \leq 1$

$$(\Omega, \Sigma, \mu) \to [0,1],\ \omega \mapsto \mathcal{E}(\delta_X)(a)$$

is a classical random variable. Applying the Central Limit Theorem to the sequence of i.i.d. random variables $\Big(\mathcal{E}(\delta_{X_i})(a)\Big)_{i\in\mathbb{N}_0}$, $\mathcal{E}(\delta_{X_0})(a) = \mathcal{E}(\delta_X)(a)$, gives

$$\lim_{n\to\infty} \hat{\nu}\Big(|\frac{1}{n}\sum_{k=0}^{n-1}\mathcal{E}(\delta_X)(a) - \theta_\infty(a)| < \frac{\sigma t}{\sqrt{n}}\Big) = \Phi(\sqrt{n}t) \ ,$$

where σ^2 is the variance of $\mathcal{E}(\delta_X)(a)$ and $\Phi(t) = \frac{1}{2\pi}\int_{-\infty}^{t} e^{-\frac{x^2}{2}}\,dx$. Further, as $E|\mathcal{E}(\delta_X)(a)|^3 < \infty$, by the Theorem of Berry-Esseen we obtain

$$\sup_t |\mathcal{F}_n(t) - \Phi(t)| < C_0\frac{E|\mathcal{E}(\delta_X)(a)|^3}{\sigma^3\sqrt{n}} \ ,$$

where $\mathcal{F}_n(t) := \hat{\nu}\Big(|\frac{1}{n}\sum_{k=0}^{n-1}\mathcal{E}(\delta_X)(a) - \theta_\infty(a)| < \frac{\sigma t}{\sqrt{n}}\Big)$ and $\frac{1}{\sqrt{2\pi}} \leq C_0 \leq 0.8$.
Another point worth mentioning is that we could achieve much better rates of convergence in experiments by repeatedly choosing new states φ_i to decompose the transition operator (Figure 5.9). Concretely, given the decomposition $T_\psi = \alpha_0 E_{\varphi_0} + (1-\alpha_0)R_0$ we may produce n_0 random samples θ_{i_0} $i_0 = 1, \ldots, n_0$, of

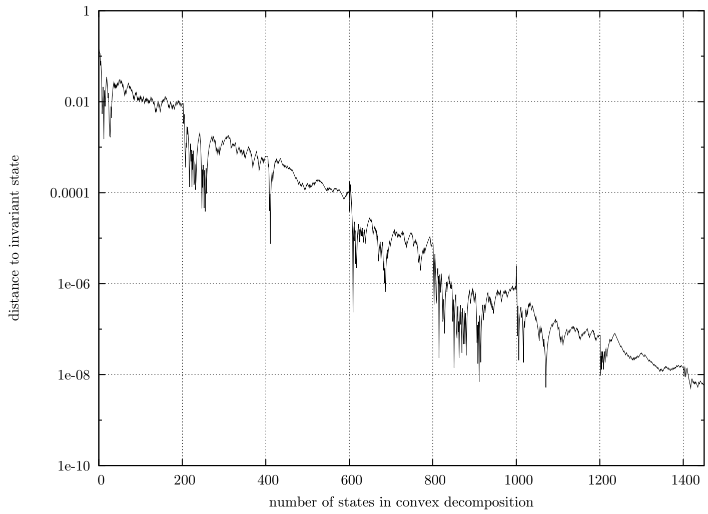


Figure 5.9: Single trajectory of example 2 with changing decomposition

the invariant state of a transition operator, determine their convex combination $\varphi_1 := \frac{1}{n_0}\sum_{i_0=1}^{n_1}\theta_{i_0}$ and decompose T_ψ into the conditional expectation E_{φ_1} and some remainder R_1:

$$T_\psi = \alpha_1 E_{\varphi_1} + (1-\alpha_1)R_1 \ .$$

With respect to that decomposition we may produce random samples θ_{i_1} $i_1 = 1,\ldots,n_1$, and again determine their convex combination $\varphi_2 := \frac{1}{n_1}\sum_{i_1=1}^{n_1}\theta_{i_1}$ and corresponding decomposition $T_\psi = \alpha_2 E_{\varphi_2} + (1-\alpha_2)R_2$. Repeating that procedure we obtained K families of random samples $\{\{\theta_{i_k}\}_{i_k}\}_{k\in K}$. Be careful however, for $\{\theta_{i_k}\}_{i_k=1,\ldots,n_k}$ is not one family of random samples. Only the random samples produced by one decomposition, i.e. the random samples θ_{i_k} $i_k = 1,\ldots,n_k$, for a fixed k, provide a family of random samples.

GLOSSARY

$\mathbb{N}$	natural numbers, 9
$\mathbb{Z}$	integer number, 9
$\mathbb{N}_0$	$\mathbb{N} \cup \{ 0 \}$, 9
$\mathbb{R}$	real numbers, 9
$\mathbb{R}^+$	$\{ x \in \mathbb{R} \colon x \geq 0 \}$, 9
$\mathbb{C}$	complex numbers, 9
$\mathcal{H},\ \mathcal{K}$	Hilbert spaces, 9
$\mathcal{B}(\mathcal{H})$	set of linear and bounded operators, 9
M_n	$n \times n$- matrices, 9
$\mathcal{A}$	C^*-, von Neumann or W^*-algebra, 10
$\mathcal{A}_*$	predual of a W^*-algebra $\mathcal{A}$, 10
$\mathcal{A}_+$	set of positive elements in a C^*-algebra $\mathcal{A}$, 10
$\|\|x\|\|_\varphi$	φ-norm, 11
$(\mathcal{A}, \varphi)$	non-commutative probability space, 12
C^+	$C^{\mathbb{N}_0}$, 21
A^+	$A \times C^+$, 21
$\mathcal{C}^+$	$\bigotimes_{n \in \mathbb{N}_0} \mathcal{C}$, 22
$\hat{C}$	$C^{\mathbb{Z}}$, 24
$\hat{C}$	$\bigotimes_{\mathbb{Z}} C$, 25
G_γ	graph asociated to mapping γ, 37

$\mathcal{M}(\Gamma)$	set of probability measures on Γ, 42
F_A	space of functions from A to A, 53
$(\hat{B}_n)_{n\in\mathbb{N}}$	backward composition on F_A, 55
$(\hat{F}_n)_{n\in\mathbb{N}}$	forward composition on F_A, 55
$CP_u(\mathcal{A},\mathcal{A})$, $CP_u(\mathcal{A})$	unital, completely positive operators on $\mathcal{A}$, 71
E_φ	conditional expectation $\mathcal{A} \ni x \mapsto \varphi(x)\mathbb{1} \in \mathcal{A}$, 72
rank_{cs}	cs-rank, 84

INDEX

aperiodic, 42
asymtotically complete, 33, 34
automorphism, 14

backward n-composition, 73

composition process, 55, 64
  backward, 55, 64
  forward, 55, 64
concrete representation, 81
conditional expectation, 16
cs-rank, 85
C^*-algebra, 10
cycle, 37

dynamical system, 15
  reversible, 15

finite algebra, 11
functional, 11
  normal, 11
  positive, 11

global coupling, 61
  synchronising, 62
graph, 35
  aperiodic, 37
  c-, 36
  directed, 35
  irreducible, 37
  period, 37
  road-coloured, 36
  stochastic, 36
  weighted, 36

isomorphism, 14

Kraus decomposition, 81
  full, 82

Markov chain, 40
  irreducible, 41
  periodic, 41
Markov process, 19
micromaser, 27–30
morphism, 14

von Neumann algebra, 10

operator

full, 82
irreducible, 43
periodic, 44
synchronisable, 72

positive element, 11

random variable, 17

state, 11
absorbing, 77
faithful, 11
pure, 11
Stinespring representation, 80
Stinespring representation
cyclic, 80
stochastic process, 18
stationary, 18
synchronising convex decomposition, 72
synchronising word, 36

time translation, 18
topology, 10
σ-strong, 11
σ-weak, 10
uniform, 10
weak, 10
transition operator, 20

update function, 66

W^*-algebra, 10

BIBLIOGRAPHY

[Acc76] L. Accardi. Nonrelativistic Qunatum Mechanics as a Noncommutative Markov Process. *Advances in Mathematics*, 1976.

[AFL82] L. Accardi, A. Frigerio, and J.T. Lewis. Quantum Stochastic Processes. *Publications of the R.I.M.S, Kyoto*, 1982.

[Arv69] W. Arveson. Subalgebras of C^*-algebras. *Acta Mathematica*, 123:141–224, 1969. 10.1007/BF02392388.

[AU82] Alberti and Uhlmann. *Stochasticity and Partial Order - Doubly Stochastic Maps and Unitary Mixing.* D.Reidel Publ. Company, 1982.

[Cho75] M.D. Choi. Completely Positive Maps. *Linear Algebra and Applications*, 1975.

[EHK78] D.E. Evans and R. Høegh-Krohn. Spectral Properties of Positive Maps on C^*-Algebras. *Journal of the London Mathematical Society*, 1978.

[FB09] F. Fagnola and R. P. Bidot. Irreducible and Periodic Positive Maps. *Communications in Stochastic Analysis*, 2009.

[FR02] F. Fagnola and R. Rebolledo. Subharmonic Projections for a Quantum Markov Semigroup. *Journal Mathematical Physics*, 2002.

[Fri78] A. Frigerio. Stationary States of Qunatum Dynamical Semigroups. *Communications in Mathematical Physics*, 1978.

[GKL06] R. Gohm, B. Kümmerer, and T. Lang. Noncommutative Symbolic Coding. *J. Ergodic Theory and Dynamical Systems*, 2006.

[Gro81] U. Groh. The Peripheral Point Spectrum of Schwarz Operators on C^*-Algebras. *Mathematische Zeitschrift*, 1981.

[Haa06] F. Haag. *Asymptotisches Verhalten von Quanten-Markov-Halbgruppen und Quanten-Markov-Prozessen.* Shaker, 2006.

[KM00] B. Kümmerer and H. Maassen. A Scattering Theory for Markov Chains. *Infinite dimensional Analysis, Quantum Probability and Related Topics*, 2000.

[KN79] B. Kümmerer and R. Nagel. Mean Ergodic Semigroups on W^*-Algebras. *Acta Scientarium Mathematicarum, Szeged*, 1979.

[KR50] M. G. Krein and M.A. Rutman. Linear Operators Leaving a Cone Invariant in a Banach Space. *American Mathematical Society Transl.*, 1950.

[Küm] B. Kümmerer. Adjoints of Operators on W^*-Algebras. Preprint.

[Lan03] T. Lang. *Ein streutheoretischer Zugang zu Kodierungsproblemen von klassischen und Quanten-Markov-Prozessen.* Shaker, 2003.

[LP67] P.D. Lax and R.S. Phillips. *Scattering Theory.* Academic Press, 1967.

[LT91] M. Ledoux and M. Talagrand. *Probability in Banachspaces.* Springer, 1991.

[Pet97] K. Petersen. *Ergodic Theory.* Cambridge University Press, 1997.

[PW96] J.G. Propp and D.B. Wilson. Exact sampling with coupled markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 1996.

[Sak97] S. Sakai. *C^*-algebras and W^*-algebras.* Springer, 1997.

[Sen81] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer, 1981.

[SNF70] B. Sz.-Nagy and C. Foias. *Harmonic Analysis of Operators on Hilbert space*. North Holland Publishing Company, 1970.

[Sti55] W. F. Stinespring. Positive Functions on C^*-algebras. *Proceedings of the American Mathematical Society*, 1955.

[Tak03a] M. Takesaki. *Theory of Operator Algebras I*. Springer, 2003.

[Tak03b] M. Takesaki. *Theory of Operator Algebras II*. Springer, 2003.

[Tra09] A. N. Trahtman. The Road Coloring Problem. *Israel Journal of Mathematics*, 2009.

[Vea00] B.T.H. Varcoe and et al. Perparing Pure Photon Number States of the Radiation Field. *Letters to Nature*, 2000.

Kurzer Lebenslauf

06/1998	–	Abitur an der Albert Einstein Schule in Groß-Bieberau
09/1998–09/1999	–	Zivildienst beim Roten Kreuz in Erbach im Odenwald
10/1999–09/2002	–	Physikstudium an der Technischen Universität Darmstadt
10/2002–12/2003	–	Physikstudium an der University of Wales, Aberystwyth (UK)
04/2004	–	Master of Philosophy in physics
01/2005–12/2010	–	Wissenschaftlicher Mitarbeiter am Fachbereich Mathematik der Technischen Universität Darmstadt